SEQUENCE LISTING

<110> HAWIGER, DANIEL
NUSSENZWEIG, MICHEL
STEINMAN, RALPH M.
BONIFAZ, LAURA

<120> ENHANCED ANTIGEN DELIVERY AND MODULATION OF THE IMMUNE RESPONSE THEREFROM

<130> RUJ-001CNCP2

<140> 10/800,023
<141> 2004-03-12

<150> 09/925,284
<151> 2001-08-09

<150> 09/586,704
<151> 2000-06-05

<150> PCT/US96/01383
<151> 1996-01-31

<150> 08/381,528
<151> 1995-01-31

<160> 41

<170> PatentIn Ver. 3.3

<210> 1
<211> 30
<212> PRT
<213> Homo sapiens

<400> 1
Arg His Arg Leu His Leu Ala Gly Phe Ser Ser Val Arg Tyr Ala Gln
1 5 10 15

Gly Val Asn Glu Asp Glu Ile Met Leu Pro Ser Phe His Asp
20 25 30

<210> 2
<211> 25
<212> PRT
<213> Homo sapiens

<400> 2
Ser Glu Ser Ser Gly Asn Asp Pro Phe Thr Ile Val His Glu Asn Thr
1 5 10 15

Gly Lys Cys Ile Gln Pro Leu Phe Asp
20 25

<210> 3
<211> 1723
<212> PRT
<213> Mus musculus

<400> 3

```
Met Arg Thr Gly Arg Val Thr Pro Gly Leu Ala Ala Gly Leu Leu Leu
  1               5                  10                  15

Leu Leu Leu Arg Ser Phe Gly Leu Val Glu Pro Ser Glu Ser Ser Gly
             20                  25                  30

Asn Asp Pro Phe Thr Ile Val His Glu Asn Thr Gly Lys Cys Ile Gln
         35                  40                  45

Pro Leu Ser Asp Trp Val Val Ala Gln Asp Cys Ser Gly Thr Asn Asn
     50                  55                  60

Met Leu Trp Lys Trp Val Ser Gln His Arg Leu Phe His Leu Glu Ser
 65                  70                  75                  80

Gln Lys Cys Leu Gly Leu Asp Ile Thr Lys Ala Thr Asp Asn Leu Arg
                 85                  90                  95

Met Phe Ser Cys Asp Ser Thr Val Met Leu Trp Trp Lys Cys Glu His
            100                 105                 110

His Ser Leu Tyr Thr Ala Ala Gln Tyr Arg Leu Ala Leu Lys Asp Gly
        115                 120                 125

Tyr Ala Val Ala Asn Thr Asn Thr Ser Asp Val Trp Lys Lys Gly Gly
    130                 135                 140

Ser Glu Glu Asn Leu Cys Ala Gln Pro Tyr His Glu Ile Tyr Thr Arg
145                 150                 155                 160

Asp Gly Asn Ser Tyr Gly Arg Pro Cys Glu Phe Pro Phe Leu Ile Gly
                165                 170                 175

Glu Thr Trp Tyr His Asp Cys Ile His Asp Glu Asp His Ser Gly Pro
            180                 185                 190

Trp Cys Ala Thr Thr Leu Ser Tyr Glu Tyr Asp Gln Lys Trp Gly Ile
        195                 200                 205

Cys Leu Leu Pro Glu Ser Gly Cys Glu Gly Asn Trp Glu Lys Asn Glu
    210                 215                 220

Gln Ile Gly Ser Cys Tyr Gln Phe Asn Asn Gln Glu Ile Leu Ser Trp
225                 230                 235                 240

Lys Glu Ala Tyr Val Ser Cys Gln Asn Gln Gly Ala Asp Leu Leu Ser
                245                 250                 255

Ile His Ser Ala Ala Glu Leu Ala Tyr Ile Thr Gly Lys Glu Asp Ile
            260                 265                 270

Ala Arg Leu Val Trp Leu Gly Leu Asn Gln Leu Tyr Ser Ala Arg Gly
        275                 280                 285

Trp Glu Trp Ser Asp Phe Arg Pro Leu Lys Phe Leu Asn Trp Asp Pro
    290                 295                 300

Gly Thr Pro Val Ala Pro Val Ile Gly Gly Ser Ser Cys Ala Arg Met
305                 310                 315                 320

Asp Thr Glu Ser Gly Leu Trp Gln Ser Val Ser Cys Glu Ser Gln Gln
                325                 330                 335
```

```
Pro Tyr Val Cys Lys Lys Pro Leu Asn Asn Thr Leu Glu Leu Pro Asp
            340                 345                 350

Val Trp Thr Tyr Thr Asp Thr His Cys His Val Gly Trp Leu Pro Asn
        355                 360                 365

Asn Gly Phe Cys Tyr Leu Leu Ala Asn Glu Ser Ser Ser Trp Asp Ala
    370                 375                 380

Ala His Leu Lys Cys Lys Ala Phe Gly Ala Asp Leu Ile Ser Met His
385                 390                 395                 400

Ser Leu Ala Asp Val Glu Val Val Val Thr Lys Leu His Asn Gly Asp
                405                 410                 415

Val Lys Lys Glu Ile Trp Thr Gly Leu Lys Asn Thr Asn Ser Pro Ala
            420                 425                 430

Leu Phe Gln Trp Ser Asp Gly Thr Glu Val Thr Leu Thr Tyr Trp Asn
        435                 440                 445

Glu Asn Glu Pro Ser Val Pro Phe Asn Lys Thr Pro Asn Cys Val Ser
    450                 455                 460

Tyr Leu Gly Lys Leu Gly Gln Trp Lys Val Gln Ser Cys Glu Lys Lys
465                 470                 475                 480

Leu Arg Tyr Val Cys Lys Lys Lys Gly Glu Ile Thr Lys Asp Ala Glu
                485                 490                 495

Ser Asp Lys Leu Cys Pro Pro Asp Glu Gly Trp Lys Arg His Gly Glu
            500                 505                 510

Thr Cys Tyr Lys Ile Tyr Glu Lys Glu Ala Pro Phe Gly Thr Asn Cys
        515                 520                 525

Asn Leu Thr Ile Thr Ser Arg Phe Glu Gln Glu Phe Leu Asn Tyr Met
    530                 535                 540

Met Lys Asn Tyr Asp Lys Ser Leu Arg Lys Tyr Phe Trp Thr Gly Leu
545                 550                 555                 560

Arg Asp Pro Asp Ser Arg Gly Glu Tyr Ser Trp Ala Val Ala Gln Gly
                565                 570                 575

Val Lys Gln Ala Val Thr Phe Ser Asn Trp Asn Phe Leu Glu Pro Ala
            580                 585                 590

Ser Pro Gly Gly Cys Val Ala Met Ser Thr Gly Lys Thr Leu Gly Lys
        595                 600                 605

Trp Glu Val Lys Asn Cys Arg Ser Phe Arg Ala Leu Ser Ile Cys Lys
    610                 615                 620

Lys Val Ser Glu Pro Gln Glu Pro Glu Glu Ala Ala Pro Lys Pro Asp
625                 630                 635                 640

Asp Pro Cys Pro Glu Gly Trp His Thr Phe Pro Ser Ser Leu Ser Cys
                645                 650                 655

Tyr Lys Val Phe His Ile Glu Arg Ile Val Arg Lys Arg Asn Trp Glu
            660                 665                 670
```

```
Glu Ala Glu Arg Phe Cys Gln Ala Leu Gly Ala His Leu Pro Ser Phe
        675                 680                 685

Ser Arg Arg Glu Glu Ile Lys Asp Phe Val His Leu Leu Lys Asp Gln
    690                 695                 700

Phe Ser Gly Gln Arg Trp Leu Trp Ile Gly Leu Asn Lys Arg Ser Pro
705                 710                 715                 720

Asp Leu Gln Gly Ser Trp Gln Trp Ser Asp Arg Thr Pro Val Ser Ala
                725                 730                 735

Val Met Met Glu Pro Glu Phe Gln Gln Asp Phe Asp Ile Arg Asp Cys
            740                 745                 750

Ala Ala Ile Lys Val Leu Asp Val Pro Trp Arg Arg Val Trp His Leu
        755                 760                 765

Tyr Glu Asp Lys Asp Tyr Ala Tyr Trp Lys Pro Phe Ala Cys Asp Ala
    770                 775                 780

Lys Leu Glu Trp Val Cys Gln Ile Pro Lys Gly Ser Thr Pro Gln Met
785                 790                 795                 800

Pro Asp Trp Tyr Asn Pro Glu Arg Thr Gly Ile His Gly Pro Pro Val
                805                 810                 815

Ile Ile Glu Gly Ser Glu Tyr Trp Phe Val Ala Asp Pro His Leu Asn
            820                 825                 830

Tyr Glu Glu Ala Val Leu Tyr Cys Ala Ser Asn His Ser Phe Leu Ala
        835                 840                 845

Thr Ile Thr Ser Phe Thr Gly Leu Lys Ala Ile Lys Asn Lys Leu Ala
    850                 855                 860

Asn Ile Ser Gly Glu Glu Gln Lys Trp Trp Val Lys Thr Ser Glu Asn
865                 870                 875                 880

Pro Ile Asp Arg Tyr Phe Leu Gly Ser Arg Arg Arg Leu Trp His His
                885                 890                 895

Phe Pro Met Thr Phe Gly Asp Glu Cys Leu His Met Ser Ala Lys Thr
            900                 905                 910

Trp Leu Val Asp Leu Ser Lys Arg Ala Asp Cys Asn Ala Lys Leu Pro
        915                 920                 925

Phe Ile Cys Glu Arg Tyr Asn Val Ser Ser Leu Glu Lys Tyr Ser Pro
    930                 935                 940

Asp Pro Ala Ala Lys Val Gln Cys Thr Glu Lys Trp Ile Pro Phe Gln
945                 950                 955                 960

Asn Lys Cys Phe Leu Lys Val Asn Ser Gly Pro Val Thr Phe Ser Gln
                965                 970                 975

Ala Ser Gly Ile Cys His Ser Tyr Gly Gly Thr Leu Pro Ser Val Leu
            980                 985                 990

Ser Arg Gly Glu Gln Asp Phe Ile Ile Ser Leu Leu Pro Glu Met Glu
        995                1000                1005
```

```
Ala Ser Leu Trp Ile Gly Leu Arg Trp Thr Ala Tyr Glu Arg Ile Asn
   1010                1015                1020

Arg Trp Thr Asp Asn Arg Glu Leu Thr Tyr Ser Asn Phe His Pro Leu
1025                1030                1035                1040

Leu Val Gly Arg Arg Leu Ser Ile Pro Thr Asn Phe Phe Asp Asp Glu
                1045                1050                1055

Ser His Phe His Cys Ala Leu Ile Leu Asn Leu Lys Lys Ser Pro Leu
            1060                1065                1070

Thr Gly Thr Trp Asn Phe Thr Ser Cys Ser Glu Arg His Ser Leu Ser
        1075                1080                1085

Leu Cys Gln Lys Tyr Ser Glu Thr Glu Asp Gly Gln Pro Trp Glu Asn
   1090                1095                1100

Thr Ser Lys Thr Val Lys Tyr Leu Asn Asn Leu Tyr Lys Ile Ile Ser
1105                1110                1115                1120

Lys Pro Leu Thr Trp His Gly Ala Leu Lys Glu Cys Met Lys Glu Lys
                1125                1130                1135

Met Arg Leu Val Ser Ile Thr Asp Pro Tyr Gln Gln Ala Phe Leu Ala
            1140                1145                1150

Val Gln Ala Thr Leu Arg Asn Ser Ser Phe Trp Ile Gly Leu Ser Ser
        1155                1160                1165

Gln Asp Asp Glu Leu Asn Phe Gly Trp Ser Asp Gly Lys Arg Leu Gln
   1170                1175                1180

Phe Ser Asn Trp Ala Gly Ser Asn Glu Gln Leu Asp Asp Cys Val Ile
1185                1190                1195                1200

Leu Asp Thr Asp Gly Phe Trp Lys Thr Ala Asp Cys Asp Asp Asn Gln
                1205                1210                1215

Pro Gly Ala Ile Cys Tyr Tyr Pro Gly Asn Glu Thr Glu Glu Glu Val
            1220                1225                1230

Arg Ala Leu Asp Thr Ala Lys Cys Pro Ser Pro Val Gln Ser Thr Pro
        1235                1240                1245

Trp Ile Pro Phe Gln Asn Ser Cys Tyr Asn Phe Met Ile Thr Asn Asn
   1250                1255                1260

Arg His Lys Thr Val Thr Pro Glu Glu Val Gln Ser Thr Cys Glu Lys
1265                1270                1275                1280

Leu His Pro Lys Ala His Ser Leu Ser Ile Arg Asn Glu Glu Glu Asn
                1285                1290                1295

Thr Phe Val Val Glu Gln Leu Leu Tyr Phe Asn Tyr Ile Ala Ser Trp
            1300                1305                1310

Val Met Leu Gly Ile Thr Tyr Glu Asn Asn Ser Leu Met Trp Phe Asp
        1315                1320                1325

Lys Thr Ala Leu Ser Tyr Thr His Trp Arg Thr Gly Arg Pro Thr Val
   1330                1335                1340
```

```
Lys Asn Gly Lys Phe Leu Ala Gly Leu Ser Thr Asp Gly Phe Trp Asp
1345                1350                1355                1360

Ile Gln Ser Phe Asn Val Ile Glu Glu Thr Leu His Phe Tyr Gln His
                1365                1370                1375

Ser Ile Ser Ala Cys Lys Ile Glu Met Val Asp Tyr Glu Asp Lys His
            1380                1385                1390

Asn Gly Thr Leu Pro Gln Phe Ile Pro Tyr Lys Asp Gly Val Tyr Ser
        1395                1400                1405

Val Ile Gln Lys Lys Val Thr Trp Tyr Glu Ala Leu Asn Ala Cys Ser
    1410                1415                1420

Gln Ser Gly Gly Glu Leu Ala Ser Val His Asn Pro Asn Gly Lys Leu
1425                1430                1435                1440

Phe Leu Glu Asp Ile Val Asn Arg Asp Gly Phe Pro Leu Trp Val Gly
                1445                1450                1455

Leu Ser Ser His Asp Gly Ser Glu Ser Ser Phe Glu Trp Ser Asp Gly
            1460                1465                1470

Arg Ala Phe Asp Tyr Val Pro Trp Gln Ser Leu Gln Ser Pro Gly Asp
        1475                1480                1485

Cys Val Val Leu Tyr Pro Lys Gly Ile Trp Arg Arg Glu Lys Cys Leu
    1490                1495                1500

Ser Val Lys Asp Gly Ala Ile Cys Tyr Lys Pro Thr Lys Asp Lys Lys
1505                1510                1515                1520

Leu Ile Phe His Val Lys Ser Ser Lys Cys Pro Val Ala Lys Arg Asp
                1525                1530                1535

Gly Pro Gln Trp Val Gln Tyr Gly Gly His Cys Tyr Ala Ser Asp Gln
            1540                1545                1550

Val Leu His Ser Phe Ser Glu Ala Lys Gln Val Cys Gln Glu Leu Asp
        1555                1560                1565

His Ser Ala Thr Val Val Thr Ile Ala Asp Glu Asn Glu Asn Lys Phe
    1570                1575                1580

Val Ser Arg Leu Met Arg Glu Asn Tyr Asn Ile Thr Met Arg Val Trp
1585                1590                1595                1600

Leu Gly Leu Ser Gln His Ser Leu Asp Gln Ser Trp Ser Trp Leu Asp
                1605                1610                1615

Gly Leu Asp Val Thr Phe Val Lys Trp Glu Asn Lys Thr Lys Asp Gly
            1620                1625                1630

Asp Gly Lys Cys Ser Ile Leu Ile Ala Ser Asn Glu Thr Trp Arg Lys
        1635                1640                1645

Val His Cys Ser Arg Gly Tyr Ala Arg Ala Val Cys Lys Ile Pro Leu
    1650                1655                1660

Ser Pro Asp Tyr Thr Gly Ile Ala Ile Leu Phe Ala Val Leu Cys Leu
1665                1670                1675                1680
```

```
Leu Gly Leu Ile Ser Leu Ala Ile Trp Phe Leu Leu Gln Arg Ser His
               1685                1690                1695

Ile Arg Trp Thr Gly Phe Ser Ser Val Arg Tyr Glu His Gly Thr Asn
           1700                1705                1710

Glu Asp Glu Val Met Leu Pro Ser Phe His Asp
       1715                1720
```

<210> 4
<211> 30
<212> PRT
<213> Mus musculus

<400> 4

```
Arg Ser His Ile Arg Trp Thr Gly Phe Ser Ser Val Arg Tyr Glu His
  1               5                  10                  15

Gly Thr Asn Glu Asp Glu Val Met Leu Pro Ser Phe His Asp
             20                  25                  30
```

<210> 5
<211> 5477
<212> DNA
<213> Homo sapiens

<400> 5

```
gaattccggg ggcgggagcc gcgtgcgccc gaggacccgg ccggaaggct tgcgccagct 60
caggatgagg acaggctggg cgacccctcg ccgcccggcg gggctcctca tgctgctctt 120
ctggttcttc gatctcgcgg agccctctgg ccgcgcagct aatgacccct tcaccatcgt 180
ccatggaaat acgggcaagt gcatcaagcc agtgtatggc tggatagtag cagacgactg 240
tgatgaaact gaggacaagt tatggaagtg ggtgtcccag catcggctct ttcatttgca 300
ctcccaaaag tgccttggcc tcgatattac caaatcggta aatgagctga gaatgttcag 360
ctgtgactcc agtgccatgc tgtggtggaa atgtgagcac cactctctgt acggagctgc 420
ccggtaccgg ctggctctga aggatggaca tggcacagca atctcaaatg catctgatgt 480
ctggaagaaa ggaggctcag aggaaagcct ttgtgaccag ccttatcatg agatctatac 540
cagagatggg aactcttatg ggagaccttg tgaatttcca ttcttaattg atgggacctg 600
gcatcatgat tgcattcttg atgaagatca tagtgggcca tggtgtgcca ccaccttaaa 660
ttatgaatat gaccgaaagt ggggcatctg cttaaagcct gaaaacggtt gtgaagataa 720
ttgggaaaag aacgagcagt ttggaagttg ctaccaattt aatactcaga cggctctttc 780
ttggaaagaa gcttatgttt catgtcagaa tcaaggagct gatttactga gcatcaacag 840
tgctgctgaa ttaacttacc ttaaagataa agaaggcatt gctaagattt tctggattgg 900
tttaaatcag ctatactctg ctagaggctg ggaatggtca gaccacaaac cattaaactt 960
tctcaactgg gatccagaca ggcccagtgc acctactata ggtggctcca gctgtgcaag 1020
aatggatgct gagtctggtc tgtggcagag cttttcctgt gaagctcaac tgccctatgt 1080
ctgcaggaaa ccattaaata atacagtgga gttaacagat gtctggacat actcagatac 1140
ccgctgtgat gcaggctggc tgccaaataa tggattttgc tatctgctgg taaatgaaag 1200
taattcctgg gataaggcac atgcgaaatg caaagccttc agtagtgacc taatcagcat 1260
tcattctcta gcagatgtgg aggtggttgt cacaaaactc cataatgagg atatcaaaga 1320
agaagtgtgg ataggcctta agaacataaa cataccaact ttatttcagt ggtcagatgg 1380
tactgaagtt actctaacat attgggatga gaatgagcca aatgttccct acaataagac 1440
gcccaactgt gtttcctact taggagagct aggtcagtgg aaagtccaat catgtgagga 1500
gaaactaaaa tatgtatgca agagaaaggg agaaaaactg aatgacgcaa gttctgataa 1560
gatgtgtcct ccagatgagg gctggaagag acatggagaa acctgttaca agatttatga 1620
ggatgaggtc ccttttggaa caaactgcaa tctgactatc actagcagat ttgagcaaga 1680
atacctaaat gatttgatga aaaagtatga taaatctcta agaaaatact tctggactgg 1740
cctgagagat gtagattctt gtggagagta taactgggca actgttggtg gaagaaggcg 1800
ggctgtaacc ttttccaact ggaattttct tgagccagct tccccgggcg gctgcgtggc 1860
tatgtctact ggaaagtctg ttggaaagtg ggaggtgaag gactgcagaa gcttcaaagc 1920
actttcaatt tgcaagaaaa tgagtggacc ccttgggcct gaagaagcat cccctaagcc 1980
tgatgacccc tgtcctgaag gctggcagag tttccccgca agtctttctt gttataaggt 2040
```

attccatgca gaaagaattg taagaaagag gaactgggaa gaagctgaac gattctgcca 2100
agcccttgga gcacaccttt ctagcttcag ccatgtggat gaaataaagg aatttcttca 2160
ctttttaacg gaccagttca gtggccagca ttggctgtgg attggtttga ataaaaggag 2220
cccagattta caaggatcct ggcaatggag tgatcgtaca ccagtgtcta ctattatcat 2280
gccaaatgag tttcagcagg attatgacat cagagactgt gctgctgtca aggtatttca 2340
taggccatgg cgaagaggct ggcatttcta tgatgataga gaatttattt atttgaggcc 2400
ttttgcttgt gatacaaaac ttgaatgggt gtgccaaatt ccaaaaggcc gtactccaaa 2460
aacaccagac tggtacaatc cagagcgtgc tggaattcat ggacctccac ttataattga 2520
aggaagtgaa tattggtttg ttgctgatct tcacctaaac tatgaagaag ccgtcctgta 2580
ctgtgccagc aatcacagct ttcttgcaac tataacatct tttgtgggac taaaagccat 2640
caaaaacaaa atagcaaata tatctggtga tggacagaag tggtggataa gaattagcga 2700
gtggccaata gatgatcatt ttacatactc acgatatcca tggcaccgct ttcctgtgac 2760
atttggagag gaatgcttgt acatgtctgc caagacttgg cttatcgact taggtaaacc 2820
aacagactgt agtaccaagt tgcccttcat ctgtgaaaaa tataatgttt cttcgttaga 2880
gaaatacagc ccagattctg cagctaaagt gcaatgttct gagcaatgga ttccttttca 2940
gaataagtgt tttctaaaga tcaaacccgt gtctctcaca ttttctcaag caagcgatac 3000
ctgtcactcc tatggtggca cccttccttc agtgttgagc cagattgaac aagactttat 3060
tacatccttg cttccggata tggaagctac tttatggatt ggtttgcgct ggactgccta 3120
tgaaaagata aacaaatgga cagataacag agagctgacg tacagtaact ttcacccatt 3180
attggttagt gggaggctga gaataccaga aaatttttt gaggaagagt ctcgctacca 3240
ctgtgcccta atactcaacc tccaaaaatc accgtttact gggacgtgga attttacatc 3300
ctgcagtgaa cgccactttg tgtctctctg tcagaaatat tcagaagtta aaagcagaca 3360
gacgttgcag aatgcttcag aaactgtaaa gtatctaaat aatctgtaca aaataatccc 3420
aaagactctg acttggcaca gtgctaaaag ggagtgtctg aaaagtaaca tgcagctggt 3480
gagcatcacg gacccttacc agcaggcatt cctcagtgtg caggcgctcc ttcacaactc 3540
ttccttatgg atcggactct tcagtcaaga tgatgaactc aactttggtt ggtcagatgg 3600
gaaacgtctt cattttagtc gctgggctga aactaatggg caactcgaag actgtgtagt 3660
attagacact gatggattct ggaaaacagt tgattgcaat gacaatcaac caggtgctat 3720
ttgctactat ccaggaaatg agactgaaaa agaggtcaaa ccagttgaca gtgttaaatg 3780
tccatctcct gttctaaata ctccgtggat accatttcag aactgttgct acaatttcat 3840
aataacaaag aataggcata tggcaacaac acaggatgaa gttcatacta aatgccagaa 3900
actgaatcca aaatcacata ttctgagtat tcgagatgaa aaggagaata actttgttct 3960
tgagcaactg ctgtacttca attatatggc ttcatgggtc atgttaggaa taacttatag 4020
aaataattct cttatgtggt ttgataagac cccactgtca tatacacatt ggagagcagg 4080
aagaccaact ataaaaaatg agaggttttt ggctggttta agtactgacg gcttctggga 4140
tattcaaacc tttaaagtta ttgaagaagc agtttatttt caccagcaca gcattcttgc 4200
ttgtaaaatt gaaatggttg actacaaaga agaatataat actacactgc cacagtttat 4260
gccatatgaa gatggtattt acagtgttat tcaaaaaaag gtaacatggt atgaagcatt 4320
aaacatgtgt tctcaaagtg gaggtcactt ggcaagcgtt cacaaccaaa atggccagct 4380
ctttctggaa gatattgtaa aacgtgatgg atttccacta tgggttgggc tctcaagtca 4440
tgatggaagt gaatcaagtt ttgaatggtc tgatggtagt acatttgact atatcccatg 4500
gaaaggccaa acatctcctg gaaattgtgt tctcttggat ccaaaaggaa cttggaaaca 4560
tgaaaaatgc aactctgtta aggatggtgc tatttgttat aaacctacaa aagctaaaaa 4620
gctgtcccgt cttacatatt catcaagatg tccagcagca aaagagaatg ggtcacggtg 4680
gatccagtac aagggtcact gttacaagtc tgatcaggca ttgcacagtt tttcagaggc 4740
caaaaaattg tgttcaaaac atgatcactc tgcaactatc gtttccataa aagatgaaga 4800
tgagaataaa tttgtgagca gactgatgag ggaaaataat aacattacca tgagagtttg 4860
gcttggatta tctcaacatt ctgttgacca gtcttggagt tggttagatg gatcagaagt 4920
gacatttgtc aaatgggaaa ataaaagtaa gagtggtgtt ggaagatgta gcatgttgat 4980
agcttcaaat gaaacttgga aaaaagttga atgtgaacat ggttttggaa gagttgtctg 5040
caaagtgcct ctgggccctg attacacagc aatagctatc atagttgcca cactaagtat 5100
cttagttctc atgggcggac tgatttggtt cctcttccaa aggcaccgtt tgcacctggc 5160
gggtttctca tcagttcgat atgcacaagg agtgaatgaa gatgagatta tgcttccttc 5220
tttccatgac taaattcttc taaaagtttt ctaatttgca ctaatgtgtt atgagaaatt 5280
agtcacttaa aatgtccagt gtcagtattt actctgctcc aaagtagaac tcttaaatac 5340
tttttcagtt gtttagatct aggcatgtgc tggtatccac agttaattcc ctgctaaatg 5400
ccatgtttat caccctaatt aatagaatgg aggggactcc aaagctggaa ctgaagtcaa 5460
attgtttgac agtaata 5477

<210> 6
<211> 1825
<212> PRT

<213> Homo sapiens

<220>
<221> MOD_RES
<222> (1744)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (1751)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (1763)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (1785)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (1787)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (1795)
<223> Variable amino acid

<220>
<221> MOD_RES
<222> (1807)..(1808)
<223> Variable amino acid

<400> 6

```
Asn Ser Gly Gly Gly Ser Arg Val Arg Pro Arg Thr Arg Pro Glu Gly
  1               5                  10                  15

Leu Arg Gln Leu Arg Met Arg Thr Gly Trp Ala Thr Pro Arg Arg Pro
             20                  25                  30

Ala Gly Leu Leu Met Leu Leu Phe Trp Phe Phe Asp Leu Ala Glu Pro
         35                  40                  45

Ser Gly Arg Ala Ala Asn Asp Pro Phe Thr Ile Val His Gly Asn Thr
     50                  55                  60

Gly Lys Cys Ile Lys Pro Val Tyr Gly Trp Ile Val Ala Asp Asp Cys
 65                  70                  75                  80

Asp Glu Thr Glu Asp Lys Leu Trp Lys Trp Val Ser Gln His Arg Leu
                 85                  90                  95

Phe His Leu His Ser Gln Lys Cys Leu Gly Leu Asp Ile Thr Lys Ser
            100                 105                 110

Val Asn Glu Leu Arg Met Phe Ser Cys Asp Ser Ser Ala Met Leu Trp
        115                 120                 125

Trp Lys Cys Glu His His Ser Leu Tyr Gly Ala Ala Arg Tyr Arg Leu
```

```
    130                 135                 140

Ala Leu Lys Asp Gly His Gly Thr Ala Ile Ser Asn Ala Ser Asp Val
145                 150                 155                 160

Trp Lys Lys Gly Gly Ser Glu Glu Ser Leu Cys Asp Gln Pro Tyr His
                165                 170                 175

Glu Ile Tyr Thr Arg Asp Gly Asn Ser Tyr Gly Arg Pro Cys Glu Phe
            180                 185                 190

Pro Phe Leu Ile Asp Gly Thr Trp His His Asp Cys Ile Leu Asp Glu
        195                 200                 205

Asp His Ser Gly Pro Trp Cys Ala Thr Thr Leu Asn Tyr Glu Tyr Asp
    210                 215                 220

Arg Lys Trp Gly Ile Cys Leu Lys Pro Glu Asn Gly Cys Glu Asp Asn
225                 230                 235                 240

Trp Glu Lys Asn Glu Gln Phe Gly Ser Cys Tyr Gln Phe Asn Thr Gln
                245                 250                 255

Thr Ala Leu Ser Trp Lys Glu Ala Tyr Val Ser Cys Gln Asn Gln Gly
            260                 265                 270

Ala Asp Leu Leu Ser Ile Asn Ser Ala Ala Glu Leu Thr Tyr Leu Lys
        275                 280                 285

Asp Lys Glu Gly Ile Ala Lys Ile Phe Trp Ile Gly Leu Asn Gln Leu
    290                 295                 300

Tyr Ser Ala Arg Gly Trp Glu Trp Ser Asp His Lys Pro Leu Asn Phe
305                 310                 315                 320

Leu Asn Trp Asp Pro Asp Arg Pro Ser Ala Pro Thr Ile Gly Gly Ser
                325                 330                 335

Ser Cys Ala Arg Met Asp Ala Glu Ser Gly Leu Trp Gln Ser Phe Ser
            340                 345                 350

Cys Glu Ala Gln Leu Pro Tyr Val Cys Arg Lys Pro Leu Asn Asn Thr
        355                 360                 365

Val Glu Leu Thr Asp Val Trp Thr Tyr Ser Asp Thr Arg Cys Asp Ala
    370                 375                 380

Gly Trp Leu Pro Asn Asn Gly Phe Cys Tyr Leu Leu Val Asn Glu Ser
385                 390                 395                 400

Asn Ser Trp Asp Lys Ala His Ala Lys Cys Lys Ala Phe Ser Ser Asp
                405                 410                 415

Leu Ile Ser Ile His Ser Leu Ala Asp Val Glu Val Val Val Thr Lys
            420                 425                 430

Leu His Asn Glu Asp Ile Lys Glu Glu Val Trp Ile Gly Leu Lys Asn
        435                 440                 445

Ile Asn Ile Pro Thr Leu Phe Gln Trp Ser Asp Gly Thr Glu Val Thr
    450                 455                 460

Leu Thr Tyr Trp Asp Glu Asn Glu Pro Asn Val Pro Tyr Asn Lys Thr
```

```
465                 470                 475                 480

Pro Asn Cys Val Ser Tyr Leu Gly Glu Leu Gly Gln Trp Lys Val Gln
                485                 490                 495

Ser Cys Glu Glu Lys Leu Lys Tyr Val Cys Lys Arg Lys Gly Glu Lys
            500                 505                 510

Leu Asn Asp Ala Ser Ser Asp Lys Met Cys Pro Pro Asp Glu Gly Trp
        515                 520                 525

Lys Arg His Gly Glu Thr Cys Tyr Lys Ile Tyr Glu Asp Glu Val Pro
    530                 535                 540

Phe Gly Thr Asn Cys Asn Leu Thr Ile Thr Ser Arg Phe Glu Gln Glu
545                 550                 555                 560

Tyr Leu Asn Asp Leu Met Lys Lys Tyr Asp Lys Ser Leu Arg Lys Tyr
                565                 570                 575

Phe Trp Thr Gly Leu Arg Asp Val Asp Ser Cys Gly Glu Tyr Asn Trp
            580                 585                 590

Ala Thr Val Gly Gly Arg Arg Arg Ala Val Thr Phe Ser Asn Trp Asn
        595                 600                 605

Phe Leu Glu Pro Ala Ser Pro Gly Gly Cys Val Ala Met Ser Thr Gly
    610                 615                 620

Lys Ser Val Gly Lys Trp Glu Val Lys Asp Cys Arg Ser Phe Lys Ala
625                 630                 635                 640

Leu Ser Ile Cys Lys Lys Met Ser Gly Pro Leu Gly Pro Glu Glu Ala
                645                 650                 655

Ser Pro Lys Pro Asp Asp Pro Cys Pro Glu Gly Trp Gln Ser Phe Pro
            660                 665                 670

Ala Ser Leu Ser Cys Tyr Lys Val Phe His Ala Glu Arg Ile Val Arg
        675                 680                 685

Lys Arg Asn Trp Glu Glu Ala Glu Arg Phe Cys Gln Ala Leu Gly Ala
    690                 695                 700

His Leu Ser Ser Phe Ser His Val Asp Glu Ile Lys Glu Phe Leu His
705                 710                 715                 720

Phe Leu Thr Asp Gln Phe Ser Gly Gln His Trp Leu Trp Ile Gly Leu
                725                 730                 735

Asn Lys Arg Ser Pro Asp Leu Gln Gly Ser Trp Gln Trp Ser Asp Arg
            740                 745                 750

Thr Pro Val Ser Thr Ile Ile Met Pro Asn Glu Phe Gln Gln Asp Tyr
        755                 760                 765

Asp Ile Arg Asp Cys Ala Ala Val Lys Val Phe His Arg Pro Trp Arg
    770                 775                 780

Arg Gly Trp His Phe Tyr Asp Asp Arg Glu Phe Ile Tyr Leu Arg Pro
785                 790                 795                 800

Phe Ala Cys Asp Thr Lys Leu Glu Trp Val Cys Gln Ile Pro Lys Gly
```

```
                805                 810                 815

Arg Thr Pro Lys Thr Pro Asp Trp Tyr Asn Pro Glu Arg Ala Gly Ile
            820                 825                 830

His Gly Pro Pro Leu Ile Ile Glu Gly Ser Glu Tyr Trp Phe Val Ala
        835                 840                 845

Asp Leu His Leu Asn Tyr Glu Glu Ala Val Leu Tyr Cys Ala Ser Asn
    850                 855                 860

His Ser Phe Leu Ala Thr Ile Thr Ser Phe Val Gly Leu Lys Ala Ile
865                 870                 875                 880

Lys Asn Lys Ile Ala Asn Ile Ser Gly Asp Gly Gln Lys Trp Trp Ile
                885                 890                 895

Arg Ile Ser Glu Trp Pro Ile Asp Asp His Phe Thr Tyr Ser Arg Tyr
            900                 905                 910

Pro Trp His Arg Phe Pro Val Thr Phe Gly Glu Glu Cys Leu Tyr Met
        915                 920                 925

Ser Ala Lys Thr Trp Leu Ile Asp Leu Gly Lys Pro Thr Asp Cys Ser
    930                 935                 940

Thr Lys Leu Pro Phe Ile Cys Glu Lys Tyr Asn Val Ser Ser Leu Glu
945                 950                 955                 960

Lys Tyr Ser Pro Asp Ser Ala Ala Lys Val Gln Cys Ser Glu Gln Trp
                965                 970                 975

Ile Pro Phe Gln Asn Lys Cys Phe Leu Lys Ile Lys Pro Val Ser Leu
            980                 985                 990

Thr Phe Ser Gln Ala Ser Asp Thr Cys His Ser Tyr Gly Gly Thr Leu
        995                1000                1005

Pro Ser Val Leu Ser Gln Ile Glu Gln Asp Phe Ile Thr Ser Leu Leu
   1010                1015                1020

Pro Asp Met Glu Ala Thr Leu Trp Ile Gly Leu Arg Trp Thr Ala Tyr
1025                1030                1035                1040

Glu Lys Ile Asn Lys Trp Thr Asp Asn Arg Glu Leu Thr Tyr Ser Asn
                1045                1050                1055

Phe His Pro Leu Leu Val Ser Gly Arg Leu Arg Ile Pro Glu Asn Phe
            1060                1065                1070

Phe Glu Glu Glu Ser Arg Tyr His Cys Ala Leu Ile Leu Asn Leu Gln
        1075                1080                1085

Lys Ser Pro Phe Thr Gly Thr Trp Asn Phe Thr Ser Cys Ser Glu Arg
   1090                1095                1100

His Phe Val Ser Leu Cys Gln Lys Tyr Ser Glu Val Lys Ser Arg Gln
1105                1110                1115                1120

Thr Leu Gln Asn Ala Ser Glu Thr Val Lys Tyr Leu Asn Asn Leu Tyr
                1125                1130                1135

Lys Ile Ile Pro Lys Thr Leu Thr Trp His Ser Ala Lys Arg Glu Cys
```

```
          1140                1145                1150

Leu Lys Ser Asn Met Gln Leu Val Ser Ile Thr Asp Pro Tyr Gln Gln
       1155                1160                1165

Ala Phe Leu Ser Val Gln Ala Leu Leu His Asn Ser Ser Leu Trp Ile
   1170                1175                1180

Gly Leu Phe Ser Gln Asp Asp Glu Leu Asn Phe Gly Trp Ser Asp Gly
1185                1190                1195                1200

Lys Arg Leu His Phe Ser Arg Trp Ala Glu Thr Asn Gly Gln Leu Glu
               1205                1210                1215

Asp Cys Val Val Leu Asp Thr Asp Gly Phe Trp Lys Thr Val Asp Cys
           1220                1225                1230

Asn Asp Asn Gln Pro Gly Ala Ile Cys Tyr Tyr Pro Gly Asn Glu Thr
       1235                1240                1245

Glu Lys Glu Val Lys Pro Val Asp Ser Val Lys Cys Pro Ser Pro Val
   1250                1255                1260

Leu Asn Thr Pro Trp Ile Pro Phe Gln Asn Cys Cys Tyr Asn Phe Ile
1265                1270                1275                1280

Ile Thr Lys Asn Arg His Met Ala Thr Thr Gln Asp Glu Val His Thr
               1285                1290                1295

Lys Cys Gln Lys Leu Asn Pro Lys Ser His Ile Leu Ser Ile Arg Asp
           1300                1305                1310

Glu Lys Glu Asn Asn Phe Val Leu Glu Gln Leu Leu Tyr Phe Asn Tyr
       1315                1320                1325

Met Ala Ser Trp Val Met Leu Gly Ile Thr Tyr Arg Asn Asn Ser Leu
   1330                1335                1340

Met Trp Phe Asp Lys Thr Pro Leu Ser Tyr Thr His Trp Arg Ala Gly
1345                1350                1355                1360

Arg Pro Thr Ile Lys Asn Glu Arg Phe Leu Ala Gly Leu Ser Thr Asp
               1365                1370                1375

Gly Phe Trp Asp Ile Gln Thr Phe Lys Val Ile Glu Glu Ala Val Tyr
           1380                1385                1390

Phe His Gln His Ser Ile Leu Ala Cys Lys Ile Glu Met Val Asp Tyr
       1395                1400                1405

Lys Glu Glu Tyr Asn Thr Thr Leu Pro Gln Phe Met Pro Tyr Glu Asp
   1410                1415                1420

Gly Ile Tyr Ser Val Ile Gln Lys Lys Val Thr Trp Tyr Glu Ala Leu
1425                1430                1435                1440

Asn Met Cys Ser Gln Ser Gly Gly His Leu Ala Ser Val His Asn Gln
               1445                1450                1455

Asn Gly Gln Leu Phe Leu Glu Asp Ile Val Lys Arg Asp Gly Phe Pro
           1460                1465                1470

Leu Trp Val Gly Leu Ser Ser His Asp Gly Ser Glu Ser Ser Phe Glu
```

```
       1475                1480                1485

Trp Ser Asp Gly Ser Thr Phe Asp Tyr Ile Pro Trp Lys Gly Gln Thr
   1490                1495                1500

Ser Pro Gly Asn Cys Val Leu Leu Asp Pro Lys Gly Thr Trp Lys His
1505                1510                1515                1520

Glu Lys Cys Asn Ser Val Lys Asp Gly Ala Ile Cys Tyr Lys Pro Thr
               1525                1530                1535

Lys Ala Lys Lys Leu Ser Arg Leu Thr Tyr Ser Ser Arg Cys Pro Ala
           1540                1545                1550

Ala Lys Glu Asn Gly Ser Arg Trp Ile Gln Tyr Lys Gly His Cys Tyr
       1555                1560                1565

Lys Ser Asp Gln Ala Leu His Ser Phe Ser Glu Ala Lys Lys Leu Cys
   1570                1575                1580

Ser Lys His Asp His Ser Ala Thr Ile Val Ser Ile Lys Asp Glu Asp
1585                1590                1595                1600

Glu Asn Lys Phe Val Ser Arg Leu Met Arg Glu Asn Asn Asn Ile Thr
               1605                1610                1615

Met Arg Val Trp Leu Gly Leu Ser Gln His Ser Val Asp Gln Ser Trp
           1620                1625                1630

Ser Trp Leu Asp Gly Ser Glu Val Thr Phe Val Lys Trp Glu Asn Lys
       1635                1640                1645

Ser Lys Ser Gly Val Gly Arg Cys Ser Met Leu Ile Ala Ser Asn Glu
   1650                1655                1660

Thr Trp Lys Lys Val Glu Cys Glu His Gly Phe Gly Arg Val Val Cys
1665                1670                1675                1680

Lys Val Pro Leu Gly Pro Asp Tyr Thr Ala Ile Ala Ile Ile Val Ala
               1685                1690                1695

Thr Leu Ser Ile Leu Val Leu Met Gly Gly Leu Ile Trp Phe Leu Phe
           1700                1705                1710

Gln Arg His Arg Leu His Leu Ala Gly Phe Ser Ser Val Arg Tyr Ala
       1715                1720                1725

Gln Gly Val Asn Glu Asp Glu Ile Met Leu Pro Ser Phe His Asp Xaa
   1730                1735                1740

Ile Leu Leu Lys Val Phe Xaa Phe Ala Leu Met Cys Tyr Glu Lys Leu
1745                1750                1755                1760

Val Thr Xaa Asn Val Gln Cys Gln Tyr Leu Leu Cys Ser Lys Val Glu
               1765                1770                1775

Leu Leu Asn Thr Phe Ser Val Val Xaa Ile Xaa Ala Cys Ala Gly Ile
           1780                1785                1790

His Ser Xaa Phe Pro Ala Lys Cys His Val Tyr His Pro Asn Xaa Xaa
       1795                1800                1805

Asn Gly Gly Asp Ser Lys Ala Gly Thr Glu Val Lys Leu Phe Asp Ser
```

1810 1815 1820

Asn
1825

<210> 7
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 7
tcgaagttgt ccttactcac aatcactcta tagcgccggc gatttgata 49

<210> 8
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 8
tcttctcaga gagggtgaga ggaccatttc gatcgatcac tcgccggcga tttgata 57

<210> 9
<211> 68
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 9
ctagcgacat ggccaagaag gagacagtct ggaggctcga ggagttcggt aggttcacaa 60
acaggaac 68

<210> 10
<211> 71
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 10
acagacggta gcacagacta tggtattctc cagattaaca gcaggtatta tgacggtagg 60
acatgatagg c 71

<210> 11
<211> 70
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 11
gtctgtgttc ctgtttgtga acctaccgaa ctcctcgagc ctccagactg tctccttctt 60
ggccatgtcg 70

<210> 12
<211> 69
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
primer

<400> 12
ggccgcctat catgtcctac cgtcataata cctgctgtta atctggagaa taccatagtc 60
tgtgctacc 69

<210> 13
<211> 330
<212> DNA
<213> Homo sapiens

<400> 13
caggctgttg tgactcagga atcagcactc accacatcac ctggtgaaac agtcacactc 60
acttgtcgct caagtactgg ggctgttaca attagtaact atgccaactg ggtccaagaa 120
aaaccagatc atttattcac tggtctaata ggtggtacca acaaccgagc tccaggtgtt 180
cctgccagat tctcaggctc cctgattgga gacaaggctg ccctcaccat cacaggggca 240
cagactgagg atgaggcaat ctatttctgt gctctatggt acaacaacca gtttatttttc 300
ggcagtggaa ccaaggtcac tgtcctaggt 330

<210> 14
<211> 354
<212> DNA
<213> Homo sapiens

<400> 14
gaggtccagc tgcaacagtc tggacctgtg ctggtgaagc ctggggcttc agtgaagatg 60
tcctgtaagg cttctggaaa cacattcact gactccttta tgcactggat gaaacagagc 120
catggaaaga gtcttgagtg gattggaatt attaatcctt ataacggcgg tactagctac 180
aaccagaaat tcaagggcaa ggccacattg actgttgaca agtcctccag cacagcctac 240
atggagctca acagcctgac atctgaggac tctgcagtct attactgtgc aagaaacggg 300
gtgcggtact actttgacta ctggggccaa ggcaccactc tcacagtctc ctca 354

<210> 15
<211> 8
<212> PRT
<213> Mus musculus

<400> 15
Ser Ile Ile Asn Phe Glu Lys Leu
1 5

<210> 16
<211> 17
<212> PRT
<213> Mus musculus

<400> 16
Leu Ser Gln Ala Val His Ala Ala His Ala Glu Ile Asn Glu Ala Gly
1 5 10 15

Arg

<210> 17
<211> 8858
<212> DNA
<213> Homo sapiens

<400> 17
gacggatcgg gagatctgct agcccgggtg acctgaggcg cgccggcttc gaatagccag 60
agtaaccttt ttttttaatt ttatttttatt ttatttttga gatggagttt ggcgccgatc 120
tcccgatccc ctatggtcga ctctcagtac aatctgctct gatgccgcat agttaagcca 180
gtatctgctc cctgcttgtg tgttggaggt cgctgagtag tgcgcgagca aaatttaagc 240
tacaacaagg caaggcttga ccgacaattg catgaagaat ctgcttaggg ttaggcgttt 300
tgcgctgctt cgcgatgtac gggccagata tacgcgttga cattgattat tgactagtta 360
ttaatagtaa tcaattacgg ggtcattagt tcatagccca tatatggagt tccgcgttac 420
ataacttacg gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc 480
aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt 540
ggactattta cggtaaactg cccacttggc agtacatcaa gtgtatcata tgccaagtac 600
gccccctatt gacgtcaatg acggtaaatg gcccgcctgg cattatgccc agtacatgac 660
cttatgggac tttcctactt ggcagtacat ctacgtatta gtcatcgcta ttaccatggt 720
gatgcggttt tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc 780
aagtctccac cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt 840
tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc gtgtacggtg 900
ggaggtctat ataagcagag ctctctggct aactagagaa cccactgctt actggcttat 960
cgaaattaat acgactcact atagggagac ccaagcttgg taccatggaa gccccagctc 1020
agcttctctt cctcctgcta ctctggctcc cagataccac cggagacatt gttctgactc 1080
agtctccagc caccctgtct gtgactccag gagatagagt ctctctttcc tgcagggcca 1140
gccagagtat tagcgactac ttacactggt atcaacaaaa atcacatgag tctccaaggc 1200
ttctcatcaa atatgcttcc cattccatct ctgggatccc ctccaggttc agtggcagtg 1260
gatcagggtc agatttcact ctcagtatca acagtgtgga acctgaagat gttggaattt 1320
attactgtca acatggtcac agctttccgt ggacgttcgg tggaggcacc aagctggaaa 1380
tcaaacgtaa gtctcgagtc tctagataac cggtcaatcg gtcaatcgat tggaattcta 1440
aactctgagg gggtcggatg acgtggccat tctttgccta aagcattgag tttactgcaa 1500
ggtcagaaaa gcatgcaaag ccctcagaat ggctgcaaag agctccaaca aaacaattta 1560
gaactttatt aaggaatagg gggaagctag gaagaaactc aaaacatcaa gattttaaat 1620
acgcttcttg gtctccttgc tataattatc tgggataagc atgctgtttt ctgtctgtcc 1680
ctaacatgcc cttatccgca aacaacacac ccaagggcag aactttgtta cttaaacacc 1740
atcctgtttg cttctttcct caggaactgt ggctgcacca tctgtcttca tcttcccgcc 1800
atctgatgag cagttgaaat ctggaactgc ctctgttgtg tgcctgctga ataacttcta 1860
tcccagagag gccaaagtac agtggaaggt ggataacgcc ctccaatcgg gtaactccca 1920
ggagagtgtc acagagcagg acagcaagga cagcacctac agcctcagca gcaccctgac 1980
gctgagcaaa gcagactacg agaaacacaa agtctacgcc tgcgaagtca cccatcaggg 2040
cctgagctcg cccgtcacaa agagcttcaa caggggagag tgttagaggg agaagtgccc 2100
ccacctgctc ctcagttcca gcctgacccc ctcccatcct ttggcctctg acccttttc 2160
cacaggggac ctaccctat tgcggtcctc agctcatct ttcacctcac ccccctcctc 2220
ctccttggct ttaattatgc taatgttgga ggagaatgaa taaataaagt gaatctttgc 2280
acctgtggtt tctctctttc ctcatttaat aattattatc tgttgtttta ccaactactc 2340
aatttctctt ataagggact aaatatgtag tcatcctaag gcacgtaacc atttataaaa 2400
atcatccttc attctatttt accctatcat cctctgcaag acagtcctcc ctcaaaccca 2460
caagccttct gtcctcacag tcccctgggc catggtagga gagacttgct tccttgtttt 2520
cccctcctca gcaagccctc atagtccttt ttaagggtga caggtcttac agtcatatat 2580
cctttgattc aattccctga gaatcaacca aagcaaattt ttcaaaagaa gaaacctgct 2640

ataaagagaa tcattcattg caacatgata taaaataaca acacaataaa agcaattaaa 2700
taaacaaaca atagggaaat gtttaagttc atcatggtac ttagacttaa tggaatgtca 2760
tgccttattt acatttttaa acaggtactg agggactcct gtctgccaag ggccgtattg 2820
agtactttcc acaacctaat ttaatccaca ctatactgtg agattaaaaa cattcattaa 2880
aatgttgcaa aggttctata aagctgagag acaaatatat tctataactc agcaatccca 2940
cttctagatg actgagtgtc cccacccacc aaaaaactat gcaagaatgt tcaaagcagc 3000
tttatttaca aaagccaaaa attggaaata gcccgattgt ccaacaatag aatgagttat 3060
taaactgtgg tatgtttata cattagaata cccaatgagg agaattaaca agctacaact 3120
atacctactc acacagatga atctcataaa aataatgtta cataagagaa actcaatgca 3180
aaagatatgt tctgtatgtt ttcatccata taaagttcaa aaccaggtaa aaataaagtt 3240
agaaatttgg atggaaatta ctcttagctg ggggtgggcg agttagtgcc tgggagaaga 3300
caagaagggg cttctggggt cttggtaatg ttctgttcct cgtgtggggt tgtgcagtta 3360
tgatctgtgc actgttctgt atacacatta tgcttcaaaa taacttcaca taaagaacat 3420
cttataccca gttaatagat agaagaggaa taagtaatag gtcaagacca acgcagctgg 3480
taagtggggg cctgggatca aatagctacc tgcctaatcc tgcccwcttg agccctgaat 3540
gagtctgcct tccagggctc aaggtgctca acaaaacaac aggcctgcta ttttcctggc 3600
atctgtgccc tgtttggcta gctaggagca cacatacata gaaattaaat gaaacagacc 3660
ttcagcaagg ggacagagga cagaattaac cttgcccaga cactggaaac ccatgtatga 3720
acactcacat gtttgggaag ggggaagggc acatgtaaat gaggactctt cctcattcta 3780
tggggcactc tggccctgcc cctctcagct actcatccat ccaacacacc tttctaagta 3840
cctctctctg cctacactct gaaggggttc aggagtaact aacacagcat cccttccctc 3900
aaatgactga caatcccttt gtcctgcttt gtttttcttt ccagtcagta ctgggaaagt 3960
ggggaaggac agtcatggag aaactacata aggaagcacc ttgcccttct gcctcttgag 4020
aatgttgatg agtatcaaat ctttcaaact ttggaggttt gagtaggggt gagactcagt 4080
aatgtccctt ccaatgacat gaacttgctc actcatccct gggggccaaa ttgaacaatc 4140
aaaggcaggc ataatccagt tatgaattct tgcggccgct tgctagcttc acgtgttgga 4200
tccaaccgcg gaagggccct attctatagt gtcacctaaa tgctagagct cgctgatcag 4260
cctcgactgt gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct 4320
tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc 4380
attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg 4440
aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg gcttctgagg 4500
cggaaagaac cagctggggc tctagggggt atccccacgc gccctgtagc ggcgcattaa 4560
gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc 4620
ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccgggcct ctcaaaaaag 4680
ggaaaaaaag catgcatctc aattagtcag caaccatagt cccgccccta actccgccca 4740
tcccgccccc aactccgccc agttccgccc attctccgcc ccatggctga ctaatttttt 4800
ttatttatgc agaggccgag gccgcctcgg cctctgagct attccagaag tagtgaggag 4860
gctttttttgg aggcctaggc ttttgcaaaa agcttggaca gctcagggct gcgatttcgc 4920
gccaaacttg acggcaatcc tagcgtgaag gctggtagga ttttatcccc gctgccatca 4980
tggttcgacc attgaactgc atcgtcgccg tgtcccaaaa tatggggatt ggcaagaacg 5040
gagacctacc ctggcctccg ctcaggaacg agttcaagta cttccaaaga atgaccacaa 5100
cctcttcagt ggaaggtaaa cagaatctgg tgattatggg taggaaaacc tggttctcca 5160
ttcctgagaa gaatcgacct ttaaaggaca gaattaatat agttctcagt agagaactca 5220
aagaaccacc acgaggagct cattttcttg ccaaaagttt ggatgatgcc ttaagactta 5280
ttgaacaacc ggaattggca agtaaagtag acatggtttg gatagtcgga ggcagttctg 5340
tttaccagga agccatgaat caaccaggcc accttagact ctttgtgaca aggatcatgc 5400
aggaatttga aagtgacacg tttttcccag aaattgattt ggggaaatat aaacttctcc 5460
cagaataccc aggcgtcctc tctgaggtcc aggaggaaaa aggcatcaag tataagtttg 5520
aagtctacga gaagaaagac taacaggaag atgctttcaa gttctctgct cccctcctaa 5580
agctatgcat ttttataaga ccatgggact tttgctggct ttagatctct ttgtgaagga 5640
accttacttc tgtggtgtga cataattgga caaactacct acagagattt aaagctctaa 5700
ggtaaatata aaatttttaa gtgtataatg tgttaaacta ctgattctaa ttgtttgtgt 5760
attttagatt ccaacctatg gaactgatga atgggagcag tggtggaatg cctttaatga 5820
ggaaaacctg ttttgctcag aagaaatgcc atctagtgat gatgaggcta ctgctgactc 5880
tcaacattct actcctccaa aaaagaagag aaaggtagaa gaccccaagg actttccttc 5940
agaattgcta agtttttga gtcatgctgt gtttagtaat agaactcttg cttgctttgc 6000
tatttacacc acaaaggaaa aagctgcact gctatacaag aaaattatgg aaaaatattc 6060
tgtaaccttt ataagtaggc ataacagtta taatcataac atactgtttt ttcttactcc 6120
acacaggcat agagtgtctg ctattaataa ctatgctcaa aaattgtgta cctttagctt 6180
tttaatttgt aaaggggtta ataaggaata tttgatgtat agtgccttga ctagagatca 6240
taatcagcca taccacattt gtagaggttt tacttgcttt aaaaaacctc ccacacctcc 6300
ccctgaacct gaaacataaa atgaatgcaa ttgttgttgt taacttgttt attgcagctt 6360
ataatggtta caaataaagc aatagcatca caaatttcac aaataaagca ttttttttcac 6420

tgcattctag ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tggatcggct 6480
ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta 6540
ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat 6600
ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct 6660
gtataccgtc gacctctagc tagagcttgg cgtaatcatg gtcatagctg tttcctgtgt 6720
gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag 6780
cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 6840
tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 6900
gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg 6960
ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat 7020
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 7080
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 7140
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc 7200
cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt 7260
ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca 7320
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg 7380
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat 7440
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta 7500
cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct 7560
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 7620
aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa 7680
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 7740
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt 7800
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca 7860
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca 7920
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc 7980
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa 8040
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc 8100
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca 8160
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat 8220
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag 8280
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac 8340
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt 8400
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt 8460
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc 8520
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 8580
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 8640
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga 8700
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 8760
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 8820
ttccgcgcac atttccccga aaagtgccac ctgacgtc 8858

<210> 18
<211> 744
<212> DNA
<213> Homo sapiens

<400> 18
gaggtgcagc tgttggagtc tgggggaggc ttggtacagc ctgggggtc cctgagactc 60
tcctgtgcag cctctggatt caccttcagt gactactaca tgagctgggt ccgccaggct 120
ccagggaagg ggctggagtg ggtctcagca attagtggta gtggtgggag cacatactac 180
gcagactccc tgaagggccg gttcaccatc tccagagaca attccaagaa cacgctgtat 240
ctgcaaatga acagcctgag agccgaggac acggccgtgt attactgtgc gagagccctt 300
gttgactaca gtaacccctc cggtatggac gtctggggcc aaggtacact ggtcaccgtg 360
agcagcggtg gaggcggttc aggcggaggt ggatccggcg gtggcggatc gcagtctgtg 420
ctgactcagc caccctcagc gtctgggacc cccgggcaga gggtcaccat ctcttgttct 480
ggaagcagct ccaacatcgg aagtaatact gttaactggt atcagcagct cccaggaacg 540
gcccccaaac tcctcatcta tagggataat cagcgaccct caggggtccc tgaccgattc 600
tctggctcca agtctggcac ctcagcctcc ctggccatca gtgggctccg gtccgaggat 660
gaggctgatt attactgtgc agcatgggat gacagcctga atggttgggt gttcggcgga 720
ggaaccaagc tgacggtcct aggt 744

<210> 19
<211> 630
<212> DNA
<213> Homo sapiens

<400> 19
ggtctatata agcagagctg ggtacgtcct cacattcagt gatcagcact gaacacagac 60
ccgtcgacgg tgatcaggac tgaacagaga gaactcacca tggagtttgg gctgagctgg 120
ctttttcttg tggctatttt aaaaggtgtc cagtgtgagg tgcagctgtt ggagtctggg 180
ggaggcttgg tacagcctgg ggggtccctg agactctcct gtgcagcctc tggattcacc 240
tttagcagct atgccatgag ctgggtccgc caggctccag ggaaggggct ggagtgggtc 300
tcagctatta gtggtagtgg tggtagcaca tactacgcag actccgtgaa gggccggttc 360
accatctcca gagacaattc caagaacacg ctgtatctgc aaatgaacag cctgagagcc 420
gaggacacgg ccgtatatta ctgtgcgaaa gatggggggt actatggttc ggggagttat 480
gggtactttg actactgggg ccagggaacc ctggtcaccg tctcctcagc tagcaccaag 540
ggcccatcgg tcttcccccc ggcaccctcc tccaagagca cctctggggg cacagcggcc 600
ctgggctgcc tggtcaagga ctacttcccc 630

<210> 20
<211> 8953
<212> DNA
<213> Human immunodeficiency virus

<220>
<221> modified_base
<222> (7216)
<223> a, c, t, g, unknown or other

<400> 20
ttctctcgac gcaggactcg gcttgctgaa gtgcacacgg caagaggcga agggcggcga 60
ctggtgagta cgccaaaaat atttttgac tagcggaggc tagaaggaga gagatgggtg 120
cgagagcgtc aatattaagt gggggaaaat tagacgattg ggaaaaaatt cggttgaggc 180
caggggaaa gaaacaatat aggataaaac atctagtatg ggcaagcagg gagctggaca 240
gatttgcact taaccctggc cttctagagt cagcaaaagg ctgtcaacaa atactagtac 300
agctccaacc agctcttcag acaggaacag aagaaattaa atcattatat aatacagtag 360
caaccctcta ttgcgtacat cagaggatag agataaaaga caccaaggaa gctttagaca 420
agatagagga aattcaaaac aaaaacaaac agcagacaca gaaagcagaa actgacaaaa 480
aagacaacag tcaggtcagt caaaattatc ctatagtgca gaatctgcaa gggcaaccgg 540
tacaccaggc cttatcacct agaactttaa atgcatgggt aaaagtgata gaagagaaag 600
ccttcagccc agaagtgata cccatgtttt cagcattatc agaaggagcc accccgcaag 660
atttaaacac catgctaaac acaatagggg gacaccaagc agctatgcaa atgttaaaag 720
atactatcaa tgaggaagct gcagaatggg acagggtaca tccagtacat gcagggcctg 780
ttgcaccagg ccaggtgaga gaaccaaggg gaagtgatat agcaggaact actagtaacc 840
tccaggaaca aataggatgg atgacaggca acccaccgat cccagtagga gaaatttata 900
aaaggtggat aattctggga ctaaataaaa tagtgagaat gtatagccct gtcagcattt 960
tggatataag acaaggacca aaagaacctt tcagagacta tgtagacaga ttctttaaag 1020
ctctaagagc tgagcaagct acacaggatg taaaaaattg gatgacagat accttgttgg 1080
tccaaaatgc aaatccagat tgcaagacca ttttaaaagc attaggatca ggagctacac 1140
tagaagaaat gatgacagca tgtcagggag tgggaggacc tggtcataaa gcaagagttt 1200
tggctgaagc aatgagccaa gtgaccaata caaacataat gatgcaaaga ggtaacttta 1260
gggatcataa aagaattgtt aagtgtttca attgtggcaa acaaggacac atagcaaaaa 1320
actgcagggc ccctagaaaa aagggctgtt ggaaatgtgg aaaggaagga caccaaatga 1380
aagactgcac tgagagacag gctaatttttt tagggaagat ttggccttcc agcaaaggga 1440
ggccagggaa ttttctccag agcagaccag agccaacagc cccaccagca gagagcctcg 1500
ggttcggaga ggagatcccc tccccgaaac aggagccgaa ggacaaggaa ctgtatcctc 1560
taacttccct cagatcactc tttggcagcg acccccttgtc acaataagaa taggggggca 1620
gctaagggaa gctctattag atacaggagc agacgataca gtattagaag aaatagattt 1680
gccaggaaaa tggaaaccaa aaatgatagg gggaattgga ggttttatca aagtgagaca 1740
gtataatgag gtacccatag aaattgaggg aaaaaaggct ataggtacag tattaatagg 1800
acctacacct gtcaacataa ttggaagaaa catgttgact cagcttggtt gtactttaaa 1860
ttttccaatt agtcctattg aaactgtacc agtaaaatta aagccaggaa tggatggccc 1920

aaaaattaaa caatggccat tgacagaaga aaaaataaaa gcattaacac aaatttgtgc 1980
agaactggaa gaggagggaa aaatttcaag aattgggcct gaaaatccat ataacacccc 2040
agtatttgcc ataaagaaaa aagacagtac taaatggaga aaattagtag attttagaga 2100
actcaataaa agaactcaag acttctggga agttcagtta ggaataccac atccagcagg 2160
gttaaaaaag aaaaaatcag tcacagtact ggatgtgggg gatgcatatt tttcagtccc 2220
tttatatgaa gatttcagga agtatactgc attcactata cctagtataa acaatgagac 2280
accagggatc agatatcagt acaacgtgct accacaggga tggaaaggat caccagcaat 2340
atttcagtgt agcatgacaa aaatcttaaa acctttaga gaaagaaacc cagaaatagt 2400
tatctaccag tacatggatg acttgtatgt gggatctgac ttggaaatag aacagcatag 2460
aagaaaaata aaggagctga gggaacatct attgaagtgg ggattttaca caccagataa 2520
aaaacatcag aaagaacctc catttctttg gatgggatat gagctccatc ctgacaaatg 2580
gacagtacaa cctatacagc tgccagaaaa agaagattgg actgtcaatg atatacaaaa 2640
gttagtggga aaactaaatt gggcaagtca aatttatcca ggaattaaaa taaaggaact 2700
atgtaaactc attaggggggg ctaaagcact aacagacata gtaccattga ctagagaagc 2760
agaattggaa ctggcagaaa acaaggagat tctaaaagaa ccagtacatg gggtatatta 2820
tgacccagca agagaattaa tagcagaagt gcagaaacaa ggactggacc aatggacata 2880
tcaaatttat caggagccat ttaaaaacct gaaaacaggg aaatatgcaa aaaggaggag 2940
tgcccacact aatgatgtaa agcaattatc acaagtggtg caaaaaatag ccttggaagc 3000
catagtgata tggggaaaaa ctcctaaatt tagactaccc atacaaaagg aaacatggga 3060
gacatggtgg acagactatt ggcaggccac ctggattcct gagtgggagt ttgtcaatac 3120
ccccccctcta gtaaaattat ggtaccaatt agaaaaggaa cccataatgg gagcagaaac 3180
tttctatgta gatggggcat ctaacaggga aactaaagta ggaaaggcag ggtatgttac 3240
tgacaaagga agacagaaag taattaccct aactgacaca acaaatcaga agactgaact 3300
acaagccatt tatttagctt tacaggattc agggatagaa gtaaacatag taacagattc 3360
acaatatgca ttggggatta ttcaagcaca accagataag agtgaatcag aattagtcaa 3420
tcaaataata gaggagttaa taaagaagga aaaggtctac ctgtcgtggg taccagcaca 3480
caaaggaatt ggaggaaatg aacaagtaga taaattagtc agttctggaa tcaggaaagt 3540
gctgtttcta gatgggatag ataaagctca agaagaacat gaaaaatatc atagcaattg 3600
gagagcaatg gctagtgatt ttaatctacc acctgtagta gcaaaagaaa tagtagctag 3660
ctgtgataaa tgtcagctaa agggggaagc catgcatgga caagtagact gtagtccagg 3720
gatatggcaa ttagattgta cacatttaga aggaaaagtt atcctggtag cagttcatgt 3780
agccagtggc tatatagaag cagaagttat cccagcagaa acaggacagg aagcagcatt 3840
ttttatatta aaattagcag gcggatggcc agtaaaagca atacatacag ataatggcag 3900
caacttcacc agtggtgctg tgaaggcagc ctgttggtgg gcagatatca aacaggaatt 3960
tggaattccc tacaatcccc aaagtcaagg agtagtagaa tctatgaata aagaattaaa 4020
gaaaatcata ggacaggtaa gagaacaagc tgaacacctt aagacagcag tacagatggc 4080
agtattcata cacaatttta aaagaaaagg ggggattggg ggatacagtg caggggaaag 4140
aataatagac ataatagcaa cagacataca aactaaagaa ttacaaaaac aaatcacaaa 4200
aattcaaaat tttcgggttt attacaggga cagcagagac ccaatttgga aaggaccagc 4260
aaaactgctc tggaaaggtg aaggggcagt agtaatacaa gacaatagtg aaataaaggt 4320
agtaccaaga agaaaagcaa agatcattag agattatgga aaacagatgg caggtgatga 4380
ttgtgtggca ggtagacagg atgaggatta acacatggaa aagtttagta aagtaccata 4440
tgaatgtttc aaagaaagct agacaatggc tgtatagaca tcactatgat agccgtcatc 4500
caaaaataag ttcagaagta cacatcccac taggagaggc tagattagta gtaacaacat 4560
attggggtct gcaaacagga gaaagagatt ggcacttggg tcagggagtc tccatagaat 4620
ggaggcggaa aaggtacaga acacaagtag accctggcct ggcagaccaa ctaattcata 4680
tgcattactt tgattgtttt tcagactctg ccataaggaa ggccatatta ggacaaatag 4740
ttagccctag gtgtgactac caagcaggac ataacaaggt aggatctcta caatatctgg 4800
cattaacagc attaataaaa ccaaaaagga gaaagccacc tttgcctagt gttcagaaac 4860
tagtagagga tagatggaac aagccccaga agaccaggga ccacagagag agccatacca 4920
tgaatggaca ctagagcttt tggaggagct taaaaatgaa gctgttagac actttcctag 4980
gccatggctc catggtttag gacagtatgt ctatagcact tatggagata catgggaagg 5040
agtcgaagcc gtaataagaa tactgcaaca actattgttt attcatttca gaatcgggtg 5100
ccatcatagc agaataggca ttataccaca gagaagaggg aggaatggag ccagtagatc 5160
ctaacagaga gccctggaac catccaggaa gtcagcctaa aactgcttgt actaattgtt 5220
attgtaaaaa gtgttgctat cattgtcaag tgtgctttct acagaagggc ttaggcattt 5280
cctatggcag gaagaagcgg agacaacgac gatcagctcc tcctggcagt aagaatcatc 5340
aagatcttat accagagcag taagtaactt aattagcata tgtaatggta tctttacaaa 5400
tagtagcaat agtagcatta atagtagcat ttttccttgc aatatgtgtg tggactatag 5460
tgtatataga atataagaaa ctgttaagac aaaggaaaat agataagtta attaatagaa 5520
taagagaaag ggcagaagac agtggtaacg agagtgatgg agacacagac gagttggctg 5580
agcttgtgga gatggggcct catgatcttt ggaatgttaa tgatttgtag tgctagagaa 5640
aacttgtggg tcacagtcta ttatggggta cctgtatgga gagatgcaaa gaccacttta 5700

ttttgcgcat ctgatgctaa agcatatagt actgaaaaac ataatgtctg ggctacacat 5760
gcttgtgtac ccacagatcc aaacccacaa gagatgagtc tgccaaatgt aacagaaaat 5820
tttaacatgt ggaaaaatga catggtagac cagatgcagg aagatataat cagtgtatgg 5880
gatgaaagct taaagccatg tgtaaagata acccctctct gtgtcacttt aaattgtagc 5940
gacgtcaata gtaacaatag tacagatagt aatagtagtg caagcaacaa tagtcctgaa 6000
atcatgaaaa actgctcttt caatgtaact acagaaataa gaaataaaag gaagcaagaa 6060
tacgcgcttt tctatagaca agatgtagta ccaattaata gtgacaataa aagttatatt 6120
ctaataaact gtaatacctc agttattaaa caggcttgtc caaaggtgtc ttttcaacca 6180
attcccatac attattgtgc tccagctggt tttgcgattc taaaatgtaa taataagact 6240
ttcaatggaa caggaccatg caaaaatgtc agtacagtac aatgtacaca cggaattaag 6300
ccagtggtat caactcaact actgctaaat ggcagtgtag cagaaggaga cataataatt 6360
agatctgaaa atatctcaga caatgctaaa aacataatag tacaacttaa tgacactgta 6420
gaaattgtgt gtaccagacc taataacaat acaagaaaag gtatacacat gggaccagga 6480
caagtgctct acgcaacagg ggaaataata ggagatataa ggaaagcata ttgtaacatt 6540
agtagaaaag attggaataa cactttacgt agagtagcta aaaaactaag agaacacttt 6600
aataaaacaa tagactttac atcaccctca ggaggggaca tagaaattac aacacatagt 6660
tttaattgtg gaggagaatt tttctattgt aatacatcaa cactgttcaa tagtagttgg 6720
gatgagaata acattaagga cacaaatagt acaaatgaca acacaactat cacaatacca 6780
tgtaaaataa aacaaattgt gagaatgtgg caaagaacag gacaagcaat atatgcccct 6840
cccatcgcag gaaacattac atgcaaatca aatattacag gattattatt gacacgtgat 6900
ggaggaaaca ggaatggcag tgagaatggc actgagacct tcagacctac aggaggaaat 6960
atgaaagata attggagaag tgaattatat aaatataaag tagtagagct tgagccacta 7020
ggagtagcac ccaccaaggc aaaaagaaga gtggtggaga gagaaaaaag agcagtggga 7080
ataggagctg tgttccttgg gttcttggga acagcaggaa gcactatggg cgcagcgtca 7140
ataacgctga cggtacaggt cagacaattg ttgtctggca tagtgcaaca gcaaagcaat 7200
ttgctgaagg ctatanaagc gcaacagcat ctgttgaagc tcactgtctg gggcattaag 7260
cagctccagg caagagtcct ggctgtggaa agatacctaa aggatcaaca gctcctagga 7320
atttggggct gctctggaaa actcatctgc accactaatg tgccctggaa tgctagttgg 7380
agtaataaat cttatgagga catttgggag aacatgacct ggatacaatg ggaaagggaa 7440
attaacaatt acacaggaat aatatacagt ctaattgaag aagcacaaaa ccagcaggaa 7500
actaatgaaa aggacttatt ggcattggac aagtggacaa atttgtggaa ttggtttaac 7560
atatcaaact ggctgtggta cataaaaata ttcataatga taataggagg cttaataggt 7620
ttaagaataa tttttgctgt gcttgctata gtaaatagag ttaggcaggg atactcacct 7680
ttgtcatttc agacccttat tccaaaccca acggaagccg acaggcccgg aggaatcgaa 7740
gaaggaggtg gagagcaagg cagaaccaga tcgattcgat tagtgaacgg attcttagct 7800
cttgcctggg acgacctgcg gagcctgtgc ctcttcagtt accaccgatt gagagacttc 7860
gtcttgattg cagcgaggac tgtgggaact ctgggactca gggggtggga gatcctcaaa 7920
tatctggtga accttgtatg gtattggggg caggaactaa agaatagtgc tattagtttg 7980
cttaatacca cagcaatagc agtagctgaa ggaacagata gaatcataga aatagcacaa 8040
agagctttta gagctattct tcacatacct agaagaataa gacagggttt agaaagagct 8100
ttgctataaa atggggaaca agtggtcaaa aagttggcct caggtaaggg acagaatgag 8160
gcgagctgct cctgctccag cagcagatgg agtgggagca gtgtctcaag atttggctaa 8220
gcatggggca atcacaagca gcaatacagc agctacaaat gatgactgtg cctggctgga 8280
agcacaaaca gaggaggagg ttggatttcc agtcagacct caggtwccat taagaccaat 8340
gacatacaaa ggagcttttg atcttagctt ctttttaaaa gaaaaggggg gactggatgg 8400
gttaatttac tccaagaaaa gacaagagat ccttgatctg tgggttcata acacacaagg 8460
ttacttccct gactggcaaa actacacacc agggccaggg accabatacc cattgacatt 8520
tggatggtgc ttcaagctag taccagttga tccaagcgaa gtagaggaag ctaatgaagg 8580
agagaacaac tgcctgttac accccgcatg ccagcatgga atagaggatg aagaaagaga 8640
agtgctaaag tggaagtttg acagctccct agcacggaga cacatagccc gagagctaca 8700
tccggagttt tacaaagact gctgacaaag aagtttctag cggggacttt ccgctgggga 8760
ctttccaggg gaggtgtggc ctgggcgggg ttggggagtg gctaaccctc agatgctgca 8820
tataagcagc tgcttttcgc ttgtactggg tctctcttgt tagaccagat ctgagcctgg 8880
gagctctctg gctaactagg gaacccactg cttaagcctc aataaagctt gccttgaggg 8940
cgcatgcaag ccg 8953

<210> 21
<211> 497
<212> PRT
<213> Human immunodeficiency virus

<400> 21

```
Met Gly Ala Arg Ala Ser Ile Leu Ser Gly Gly Lys Leu Asp Asp Trp
  1               5                  10                  15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Gln Tyr Arg Ile Lys
             20                  25                  30

His Leu Val Trp Ala Ser Arg Glu Leu Asp Arg Phe Ala Leu Asn Pro
         35                  40                  45

Gly Leu Leu Glu Ser Ala Lys Gly Cys Gln Gln Ile Leu Val Gln Leu
     50                  55                  60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Ile Lys Ser Leu Tyr Asn
 65                  70                  75                  80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
                 85                  90                  95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Ile Gln Asn Lys Asn Lys
            100                 105                 110

Gln Gln Thr Gln Lys Ala Glu Thr Asp Lys Lys Asp Asn Ser Gln Val
        115                 120                 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Pro Val His
    130                 135                 140

Gln Ala Leu Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu
145                 150                 155                 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
                165                 170                 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Ile Gly
            180                 185                 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu
        195                 200                 205

Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Val Ala
    210                 215                 220

Pro Gly Gln Val Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
225                 230                 235                 240

Ser Asn Leu Gln Glu Gln Ile Gly Trp Met Thr Gly Asn Pro Pro Ile
                245                 250                 255

Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
            260                 265                 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
        275                 280                 285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Ala Leu
    290                 295                 300

Arg Ala Glu Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr
305                 310                 315                 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
                325                 330                 335
```

```
Leu Gly Ser Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
            340                 345                 350

Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
        355                 360                 365

Gln Val Thr Asn Thr Asn Ile Met Met Gln Arg Gly Asn Phe Arg Asp
    370                 375                 380

His Lys Arg Ile Val Lys Cys Phe Asn Cys Gly Lys Gln Gly His Ile
385                 390                 395                 400

Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly
                405                 410                 415

Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe
            420                 425                 430

Leu Gly Lys Ile Trp Pro Ser Ser Lys Gly Arg Pro Gly Asn Phe Leu
        435                 440                 445

Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Leu Gly Phe
    450                 455                 460

Gly Glu Glu Ile Pro Ser Pro Lys Gln Glu Pro Lys Asp Lys Glu Leu
465                 470                 475                 480

Tyr Pro Leu Thr Ser Leu Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser
                485                 490                 495

Gln
```

<210> 22
<211> 1001
<212> PRT
<213> Human immunodeficiency virus

```
<400> 22
Phe Phe Arg Glu Asp Leu Ala Phe Gln Gln Arg Glu Ala Arg Glu Phe
  1               5                  10                  15

Ser Pro Glu Gln Thr Arg Ala Asn Ser Pro Thr Ser Arg Glu Pro Arg
             20                  25                  30

Val Arg Arg Gly Asp Pro Leu Pro Glu Thr Gly Ala Glu Gly Gln Gly
         35                  40                  45

Thr Val Ser Ser Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu
     50                  55                  60

Val Thr Ile Arg Ile Gly Gly Gln Leu Arg Glu Ala Leu Leu Asp Thr
 65                  70                  75                  80

Gly Ala Asp Asp Thr Val Leu Glu Glu Ile Asp Leu Pro Gly Lys Trp
                 85                  90                  95

Lys Pro Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg Gln
            100                 105                 110

Tyr Asn Glu Val Pro Ile Glu Ile Glu Gly Lys Lys Ala Ile Gly Thr
        115                 120                 125
```

```
Val Leu Ile Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Met Leu
    130                 135                 140

Thr Gln Leu Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr
145                 150                 155                 160

Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Ile Lys Gln
                165                 170                 175

Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Thr Gln Ile Cys Ala
            180                 185                 190

Glu Leu Glu Glu Glu Gly Lys Ile Ser Arg Ile Gly Pro Glu Asn Pro
        195                 200                 205

Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp
    210                 215                 220

Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe
225                 230                 235                 240

Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys
                245                 250                 255

Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro
            260                 265                 270

Leu Tyr Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile
        275                 280                 285

Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln
    290                 295                 300

Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr Lys Ile
305                 310                 315                 320

Leu Lys Pro Phe Arg Glu Arg Asn Pro Glu Ile Val Ile Tyr Gln Tyr
                325                 330                 335

Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Glu Gln His Arg
            340                 345                 350

Arg Lys Ile Lys Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Tyr
        355                 360                 365

Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly
    370                 375                 380

Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro
385                 390                 395                 400

Glu Lys Glu Asp Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys
                405                 410                 415

Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Ile Lys Glu Leu
            420                 425                 430

Cys Lys Leu Ile Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu
        435                 440                 445

Thr Arg Glu Ala Glu Leu Glu Leu Ala Glu Asn Lys Glu Ile Leu Lys
    450                 455                 460
```

```
Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ala Arg Glu Leu Ile Ala
465                 470                 475                 480

Glu Val Gln Lys Gln Gly Leu Asp Gln Trp Thr Tyr Gln Ile Tyr Gln
                485                 490                 495

Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Arg Arg Ser
            500                 505                 510

Ala His Thr Asn Asp Val Lys Gln Leu Ser Gln Val Val Gln Lys Ile
        515                 520                 525

Ala Leu Glu Ala Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu
    530                 535                 540

Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln
545                 550                 555                 560

Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val
                565                 570                 575

Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Met Gly Ala Glu Thr
            580                 585                 590

Phe Tyr Val Asp Gly Ala Ser Asn Arg Glu Thr Lys Val Gly Lys Ala
        595                 600                 605

Gly Tyr Val Thr Asp Lys Gly Arg Gln Lys Val Ile Thr Leu Thr Asp
    610                 615                 620

Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln
625                 630                 635                 640

Asp Ser Gly Ile Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu
                645                 650                 655

Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn
            660                 665                 670

Gln Ile Ile Glu Glu Leu Ile Lys Lys Glu Lys Val Tyr Leu Ser Trp
        675                 680                 685

Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu
    690                 695                 700

Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp Lys
705                 710                 715                 720

Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala
                725                 730                 735

Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser
            740                 745                 750

Cys Asp Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp
        755                 760                 765

Cys Ser Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys
    770                 775                 780

Val Ile Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu
785                 790                 795                 800
```

```
Val Ile Pro Ala Glu Thr Gly Gln Glu Ala Ala Phe Phe Ile Leu Lys
                805                 810                 815

Leu Ala Gly Gly Trp Pro Val Lys Ala Ile His Thr Asp Asn Gly Ser
            820                 825                 830

Asn Phe Thr Ser Gly Ala Val Lys Ala Ala Cys Trp Trp Ala Asp Ile
        835                 840                 845

Lys Gln Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val
    850                 855                 860

Glu Ser Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Glu
865                 870                 875                 880

Gln Ala Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His
                885                 890                 895

Asn Phe Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg
            900                 905                 910

Ile Ile Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys
        915                 920                 925

Gln Ile Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg
    930                 935                 940

Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly
945                 950                 955                 960

Ala Val Val Ile Gln Asp Asn Ser Glu Ile Lys Val Val Pro Arg Arg
                965                 970                 975

Lys Ala Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp
            980                 985                 990

Cys Val Ala Gly Arg Gln Asp Glu Asp
        995                1000


<210> 23
<211> 101
<212> PRT
<213> Human immunodeficiency virus

<400> 23
Met Glu Pro Val Asp Pro Asn Arg Glu Pro Trp Asn His Pro Gly Ser
  1               5                  10                  15

Gln Pro Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Tyr
            20                  25                  30

His Cys Gln Val Cys Phe Leu Gln Lys Gly Leu Gly Ile Ser Tyr Gly
        35                  40                  45

Arg Lys Lys Arg Arg Gln Arg Arg Ser Ala Pro Pro Gly Ser Lys Asn
    50                  55                  60

His Gln Asp Leu Ile Pro Glu Gln Pro Leu Phe Gln Thr Gln Arg Lys
65                  70                  75                  80

Pro Thr Gly Pro Glu Glu Ser Lys Lys Glu Val Glu Ser Lys Ala Glu
```

```
                85                  90                  95

Pro Asp Arg Phe Asp
            100
```

<210> 24
<211> 116
<212> PRT
<213> Human immunodeficiency virus

<400> 24

```
Met Ala Gly Arg Ser Gly Asp Asn Asp Asp Gln Leu Leu Leu Ala Val
  1               5                  10                  15

Arg Ile Ile Lys Ile Leu Tyr Gln Ser Asn Pro Tyr Ser Lys Pro Asn
             20                  25                  30

Gly Ser Arg Gln Ala Arg Arg Asn Arg Arg Arg Arg Trp Arg Ala Arg
         35                  40                  45

Gln Asn Gln Ile Asp Ser Ile Ser Glu Arg Ile Leu Ser Ser Cys Leu
     50                  55                  60

Gly Arg Pro Ala Glu Pro Val Pro Leu Gln Leu Pro Pro Ile Glu Arg
 65                  70                  75                  80

Leu Arg Leu Asp Cys Ser Glu Asp Cys Gly Asn Ser Gly Thr Gln Gly
                 85                  90                  95

Val Gly Asp Pro Gln Ile Ser Gly Glu Pro Cys Met Val Leu Gly Ala
            100                 105                 110

Gly Thr Lys Glu
        115
```

<210> 25
<211> 850
<212> PRT
<213> Human immunodeficiency virus

<220>
<221> MOD_RES
<222> (554)
<223> Variable amino acid

<400> 25

```
Met Glu Thr Gln Thr Ser Trp Leu Ser Leu Trp Arg Trp Gly Leu Met
  1               5                  10                  15

Ile Phe Gly Met Leu Met Ile Cys Ser Ala Arg Glu Asn Leu Trp Val
             20                  25                  30

Thr Val Tyr Tyr Gly Val Pro Val Trp Arg Asp Ala Lys Thr Thr Leu
         35                  40                  45

Phe Cys Ala Ser Asp Ala Lys Ala Tyr Ser Thr Glu Lys His Asn Val
     50                  55                  60

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Met
 65                  70                  75                  80
```

```
Ser Leu Pro Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met
                 85                  90                  95

Val Asp Gln Met Gln Glu Asp Ile Ile Ser Val Trp Asp Glu Ser Leu
            100                 105                 110

Lys Pro Cys Val Lys Ile Thr Pro Leu Cys Val Thr Leu Asn Cys Ser
        115                 120                 125

Asp Val Asn Ser Asn Asn Ser Thr Asp Ser Asn Ser Ser Ala Ser Asn
    130                 135                 140

Asn Ser Pro Glu Ile Met Lys Asn Cys Ser Phe Asn Val Thr Thr Glu
145                 150                 155                 160

Ile Arg Asn Lys Arg Lys Gln Glu Tyr Ala Leu Phe Tyr Arg Gln Asp
                165                 170                 175

Val Val Pro Ile Asn Ser Asp Asn Lys Ser Tyr Ile Leu Ile Asn Cys
            180                 185                 190

Asn Thr Ser Val Ile Lys Gln Ala Cys Pro Lys Val Ser Phe Gln Pro
        195                 200                 205

Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys
    210                 215                 220

Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr
225                 230                 235                 240

Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu
                245                 250                 255

Leu Asn Gly Ser Val Ala Glu Gly Asp Ile Ile Ile Arg Ser Glu Asn
            260                 265                 270

Ile Ser Asp Asn Ala Lys Asn Ile Ile Val Gln Leu Asn Asp Thr Val
        275                 280                 285

Glu Ile Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Gly Ile His
    290                 295                 300

Met Gly Pro Gly Gln Val Leu Tyr Ala Thr Gly Glu Ile Ile Gly Asp
305                 310                 315                 320

Ile Arg Lys Ala Tyr Cys Asn Ile Ser Arg Lys Asp Trp Asn Asn Thr
                325                 330                 335

Leu Arg Arg Val Ala Lys Lys Leu Arg Glu His Phe Asn Lys Thr Ile
            340                 345                 350

Asp Phe Thr Ser Pro Ser Gly Gly Asp Ile Glu Ile Thr Thr His Ser
        355                 360                 365

Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Thr Ser Thr Leu Phe
    370                 375                 380

Asn Ser Ser Trp Asp Glu Asn Asn Ile Lys Asp Thr Asn Ser Thr Asn
385                 390                 395                 400

Asp Asn Thr Thr Ile Thr Ile Pro Cys Lys Ile Lys Gln Ile Val Arg
                405                 410                 415
```

```
Met Trp Gln Arg Thr Gly Gln Ala Ile Tyr Ala Pro Pro Ile Ala Gly
            420                 425                 430

Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp
        435                 440                 445

Gly Gly Asn Arg Asn Gly Ser Glu Asn Gly Thr Glu Thr Phe Arg Pro
    450                 455                 460

Thr Gly Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr
465                 470                 475                 480

Lys Val Val Glu Leu Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys
                485                 490                 495

Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val
            500                 505                 510

Phe Leu Gly Phe Leu Gly Thr Ala Gly Ser Thr Met Gly Ala Ala Ser
        515                 520                 525

Ile Thr Leu Thr Val Gln Val Arg Gln Leu Leu Ser Gly Ile Val Gln
    530                 535                 540

Gln Gln Ser Asn Leu Leu Lys Ala Ile Xaa Ala Gln Gln His Leu Leu
545                 550                 555                 560

Lys Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala
                565                 570                 575

Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys
            580                 585                 590

Ser Gly Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ala Ser Trp
        595                 600                 605

Ser Asn Lys Ser Tyr Glu Asp Ile Trp Glu Asn Met Thr Trp Ile Gln
    610                 615                 620

Trp Glu Arg Glu Ile Asn Asn Tyr Thr Gly Ile Ile Tyr Ser Leu Ile
625                 630                 635                 640

Glu Glu Ala Gln Asn Gln Gln Glu Thr Asn Glu Lys Asp Leu Leu Ala
                645                 650                 655

Leu Asp Lys Trp Thr Asn Leu Trp Asn Trp Phe Asn Ile Ser Asn Trp
            660                 665                 670

Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Ile Gly Gly Leu Ile Gly
        675                 680                 685

Leu Arg Ile Ile Phe Ala Val Leu Ala Ile Val Asn Arg Val Arg Gln
    690                 695                 700

Gly Tyr Ser Pro Leu Ser Phe Gln Thr Leu Ile Pro Asn Pro Thr Glu
705                 710                 715                 720

Ala Asp Arg Pro Gly Gly Ile Glu Glu Gly Gly Gly Glu Gln Gly Arg
                725                 730                 735

Thr Arg Ser Ile Arg Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp
            740                 745                 750
```

```
Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe
        755                 760                 765

Val Leu Ile Ala Ala Arg Thr Val Gly Thr Leu Gly Leu Arg Gly Trp
    770                 775                 780

Glu Ile Leu Lys Tyr Leu Val Asn Leu Val Trp Tyr Trp Gly Gln Glu
785                 790                 795                 800

Leu Lys Asn Ser Ala Ile Ser Leu Leu Asn Thr Thr Ala Ile Ala Val
                805                 810                 815

Ala Glu Gly Thr Asp Arg Ile Ile Glu Ile Ala Gln Arg Ala Phe Arg
            820                 825                 830

Ala Ile Leu His Ile Pro Arg Arg Ile Arg Gln Gly Leu Glu Arg Ala
        835                 840                 845

Leu Leu
    850


<210> 26
<211> 204
<212> PRT
<213> Human immunodeficiency virus

<220>
<221> MOD_RES
<222> (132)
<223> Variable amino acid

<400> 26
Met Gly Asn Lys Trp Ser Lys Ser Trp Pro Gln Val Arg Asp Arg Met
  1               5                  10                  15

Arg Arg Ala Ala Pro Ala Pro Ala Ala Asp Gly Val Gly Ala Val Ser
            20                  25                  30

Gln Asp Leu Ala Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala
        35                  40                  45

Thr Asn Asp Asp Cys Ala Trp Leu Glu Ala Gln Thr Glu Glu Glu Val
    50                  55                  60

Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys
65                  70                  75                  80

Gly Ala Phe Asp Leu Ser Phe Phe Leu Lys Glu Lys Gly Gly Leu Asp
                85                  90                  95

Gly Leu Ile Tyr Ser Lys Lys Arg Gln Glu Ile Leu Asp Leu Trp Val
            100                 105                 110

His Asn Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly
        115                 120                 125

Pro Gly Thr Xaa Tyr Pro Leu Thr Phe Gly Trp Cys Phe Lys Leu Val
    130                 135                 140

Pro Val Asp Pro Ser Glu Val Glu Glu Ala Asn Glu Gly Glu Asn Asn
145                 150                 155                 160
```

```
Cys Leu Leu His Pro Ala Cys Gln His Gly Ile Glu Asp Glu Glu Arg
                165                 170                 175

Glu Val Leu Lys Trp Lys Phe Asp Ser Ser Leu Ala Arg Arg His Ile
            180                 185                 190

Ala Arg Glu Leu His Pro Glu Phe Tyr Lys Asp Cys
        195                 200
```

<210> 27
<211> 158
<212> PRT
<213> Human papillomavirus

<400> 27

```
Met Ala Arg Phe Glu Asp Pro Thr Arg Arg Pro Tyr Lys Leu Pro Asp
  1               5                  10                  15

Leu Cys Thr Glu Leu Asn Thr Ser Leu Gln Asp Ile Glu Ile Thr Cys
             20                  25                  30

Val Tyr Cys Lys Thr Val Leu Glu Leu Thr Glu Val Phe Glu Phe Ala
         35                  40                  45

Phe Lys Asp Leu Phe Val Val Tyr Arg Asp Ser Ile Pro His Ala Ala
     50                  55                  60

Cys His Lys Cys Ile Asp Phe Tyr Ser Arg Ile Arg Glu Leu Arg His
 65                  70                  75                  80

Tyr Ser Asp Ser Val Tyr Gly Asp Thr Leu Glu Lys Leu Thr Asn Thr
                 85                  90                  95

Gly Leu Tyr Asn Leu Leu Ile Arg Cys Leu Arg Cys Gln Lys Pro Leu
            100                 105                 110

Asn Pro Ala Glu Lys Leu Arg His Leu Asn Glu Lys Arg Arg Phe His
        115                 120                 125

Lys Ile Ala Gly His Tyr Arg Gly Gln Cys His Ser Cys Cys Asn Arg
    130                 135                 140

Ala Arg Gln Glu Arg Leu Gln Arg Arg Arg Glu Thr Gln Val
145                 150                 155
```

<210> 28
<211> 105
<212> PRT
<213> Human papillomavirus

<400> 28

```
Met His Gly Pro Lys Ala Thr Leu Gln Asp Ile Val Leu His Leu Glu
  1               5                  10                  15

Pro Gln Asn Glu Ile Pro Val Asp Leu Leu Cys His Glu Gln Leu Ser
             20                  25                  30

Asp Ser Glu Glu Glu Asn Asp Glu Ile Asp Gly Val Asn His Gln His
         35                  40                  45

Leu Pro Ala Arg Arg Ala Glu Pro Gln Arg His Thr Met Leu Cys Met
```

```
     50                  55                  60

Cys Cys Lys Cys Glu Ala Arg Ile Glu Leu Val Val Glu Ser Ser Ala
 65                  70                  75                  80

Asp Asp Leu Arg Ala Phe Gln Gln Leu Phe Leu Lys Thr Leu Ser Phe
                 85                  90                  95

Val Cys Pro Trp Cys Ala Ser Gln Gln
            100                 105
```

<210> 29
<211> 843
<212> DNA
<213> Human papillomavirus

<400> 29
```
cggtgtatat aaaagatgtg agaaacgcac cacaatacta tggcgcgctt tgaggatcca 60
acacggcgac cctacaagct acctgatctg tgcacggaac tgaacacttc actgcaagac 120
atagaaataa cctgtgtata ttgcaagaca gtattggaac ttacagaggt atttgaattt 180
gcattcaaag atttatttgt ggtgtataga gacagtatac cgcatgctgc atgccataaa 240
tgtatagatt tttattctag aattagagaa ttaagacatt attcagactc tgtgtatgga 300
gacacattag aaaaactaac taacactggg ttatacaatt tattaataag gtgcctgcgg 360
tgccagaaac cgttgaatcc agcagaaaaa cttagacacc ttaatgaaaa acgacgattc 420
cacaaaatag ctgggcacta tagaggccag tgccattcgt gctgcaaccg agcacgacag 480
gagagactcc aacgacgcag agaaacacaa gtataatatt aagtatgcat ggacctaagg 540
caacattgca agacattgta ttgcatttag agcctcaaaa tgaaattccg gttgaccttc 600
tatgtcacga gcaattaagc gactcagagg aagaaaacga tgaaatagat ggagttaatc 660
atcaacattt accagcccga cgagccgaac cacaacgtca cacaatgttg tgtatgtgtt 720
gtaagtgtga agctagaatt gagctagtag tagaaagctc agcagacgac cttcgagcat 780
tccagcagct gtttctgaaa accctgtcct ttgtgtgtcc gtggtgtgca tcccagcagt 840
aag                                                               843
```

<210> 30
<211> 11835
<212> DNA
<213> Epstein Barr virus

<400> 30
```
ggccgctgtt cacctaaagt gacgcaaggt ctgtcagccg ccagggtccg tttaccaggc 60
tttcaggtgt ggaatttaga tagagtgggt gtgtgctctt gtttaattac accaagatca 120
ccaccctcta tccatatccc acaattgata aacctccgca tgtccaacca ccacgttgaa 180
caggatgtgg caccctaaga ggacgcaggc atacaaggtt attacccagt ccttgtatgc 240
ctggtgtccc cttagtggga cgcaggccta ggtagcatca tttacactaa aagcagtgac 300
cttgttggta ctttaaggtt ggtccaatcc ataggctttt tttgtgaaaa cccggggatc 360
ggactagcct tagagtaact caaggccaag catttcacac ctgcaaatgc accatgtaac 420
cacagatcta aactgaaagt tgcagcttta gatggcaagg aaacttgggt ttcaggcata 480
gaaagcctgg ctcactatag cagcccatgt ttgttccagg gtgggggaaa ggcacgtgcc 540
cttagaaaac ttagctgcaa aaattctatt gtgttgggag agcctctata tctaaaggcc 600
tttcctcaca atacaaatgt tactaacgtc tgccctctgg agacctgcta tgtggctaga 660
cgtatggcct acccaagacg ttgggggtct cgggtaggcc atgattcttc caggcatagg 720
ttacaaccag tcactgctat caagcctact cagttcccaa cgcagcacat acccccccgcc 780
tctcctgcca tgaggactta tggcagtgtt tactgttctg cttttactct tggaccaggc 840
tgtcattcta tcagaataac aggggaagca aggccccctg cttcagcggg acacgtgttt 900
ctagaatctc ggagccaata actacctgcc cctctaatct gtatgctgca tgaaaaacca 960
catacacgtg atgtaagttt agccagttta ttgttacacc aatgccccga aagtctcccc 1020
ctgtcccttt gggtctcagg acccagccct ggagctcggg ggcggccggg tggcccaccg 1080
ggtccgctgg gtccgctgcc ccgctccggc ggggggtggc cggctgcagc cgggtccggg 1140
gttccggccc tggagctcgg ggggcggccg ggtggcccac cgggtccgct gggtccgctg 1200
ccccgctccg gcgggggggtg gccggctgca gccgggtccg gggttccggc cctggagctc 1260
gggggggcggc cgggtggccc accgggtccg ctgggtccgc tgccccgctc cggcgggggg 1320
```

tggccggctg cagccgggtc cggggttccg gccctggagc tcgggggggcg gccgggtggc 1380
ccaccgggtc cgctgggtcc gctgccccgc tccggcgggg ggtggccggc tgcagccggg 1440
tccggggttc cggccctgga gctcgggggg cggccgggtg gcccaccggg tccgctgggt 1500
ccgctgcccc gctccggcgg ggggtggccg gctgcagccg ggtccggggt tccggccctg 1560
gagctcgggg ggcggccggg tggcccaccg ggtccgctgg gtccgctgcc ccgctccggc 1620
gggggggtggc cggctgcagc cgggtccggg gttccggccc tggagctcgg ggggcggccg 1680
ggtggcccac cgggtccgct gggtccgctg ccccgctccg gcgggggggtg gccggctgca 1740
gccgggtccg gggttccggc cctggagctc gggggggcggc cgggtggccc accgggtccg 1800
ctgggtccgc tgccccgctc cggcgggggg tggccggctg cagccgggtc cggggttccg 1860
gccctggagc tcgggggggcg gccgggtggc ccaccgggtc cgctgggtcc gctgccccgc 1920
tccggcgggg ggtggccggc tgcagccggg tccggggttc cggccctgga gctcgggggg 1980
cggccgggtg gcccaccggg tccgctgggt ccgctgcccc gctccggcgg ggggtggccg 2040
gctgcagccg ggtccggggt tccggccctg gagctcgggg ggcggccggg tggcccaccg 2100
ggtccgctgg gtccgctgcc ccgctccggc gggggggtggc cggctgcagc cgggtccggg 2160
gttccggccc tggagctcgg ggggcggccg ggtggcccac cgggtccgct gggtccgctg 2220
ccccgctccg gcgggggggtg gccggctgca gccgggtccg gggttccggc cctggagctc 2280
gggggggcggc cgggtggccc accgggtccg ctgggtccgc tgccccgctc cggcgggggg 2340
tggccggctg cagccgggtc cggggttccg gccctggagc tcgggggggcg gccgggtggc 2400
ccaccgggtc cgctgggtcc gctgccccgc tccggcgggg ggtggccggc tgcagccggg 2460
tccggggttc cggccctgga gctcgggggg cggccgggtg gcccaccggg tccgctgggt 2520
ccgctgcccc gctccggcgg ggggtggccg gctgcagccg ggtccggggt tccggccctg 2580
gagctcgggg ggcggccggg tggcccaccg ggtccgctgg gtccgctgcc ccgctccggc 2640
gggggggtggc cggctgcagc cgggtccggg gttccggccc tggagctcgg ggggcggccg 2700
ggtggcccac cgggtccgct gggtccgctg ccccgctccg gcgggggggtg gccggctgca 2760
gccgggtccg gggttccggc cctggagctc gggggggcggc cgggtggccc accgggtccg 2820
ctgggtccgc tgccccgctc cggcgggggg tggccggctg cagccgggtc cggggttccg 2880
gccctggagc tcgggggggcg gccgggtggc ccaccgggtc cgctgggtcc gctgccccgc 2940
tccggcgggg ggtggccggc tgcagccggg tccggggttc cggccctgga gctcgggggg 3000
cggccgggtg gcccaccggg tccgctgggt ccgctgcccc gctccggcgg ggggtggccg 3060
gctgcagccg ggtccggggt tccggccctg gagctcgggg ggcggccggg tggcccaccg 3120
ggtccgctgg gtccgctgcc ccgctccggc gggggggtggc cggctgcagc cgggtccggg 3180
gttccggccc tggagctcgg ggggcggccg ggtggcccac cgggtccgct gggtccgctg 3240
ccccgctccg gcgggggggtg gccggctgca gccgggtccg gggttccggc cctggagctc 3300
gggggggcggc cgggtggccc accgggtccg ctgggtccgc tgccccgctc cggcgggggg 3360
tggccggctg cagccgggtc cggggttccg gccctggagc tcgggggggcg gccgggtggc 3420
ccaccgggtc cgctgggtcc gctgccccgc tccggcgggg ggtggccggc tgcagccggg 3480
tccggggttc cggccctgga gctcgggggg cggccgggtg gcccaccggg tccgctgggt 3540
ccgctgcccc gctccggcgg ggatgggggt gcgctcccag gccggaccct ggtgccaggc 3600
agggaccccg cgccacccgc ttcatggggg gggaggccgc cgcaaggacg ccgggccggc 3660
tgggaggtgt gcacccccgg agcgtctgga cgacgctggc gagccgggcc agctcgcctt 3720
cttttatcct ctttttgggg tctctgtgca ataccttaag gtttgctcag gagtgggggg 3780
cttctcattg gttaattcag gtgtgtgatt ttagcccgtt gggttacatt aaggtgtgta 3840
accaggtggg tggtacctgg aggtcattct attgggataa cgagaggagg aggggctaga 3900
ggcccgcgag atttggggta ggcggagcct caggagggtc ccctccatag ggttgaacca 3960
ggaggggggag gatcgggctc cgccccgata tacctagtgg gtggagccta gaggtaggta 4020
tccatagggt tccattatcc tggaggtatc ctaagctccg cccctatata ccaggtgggt 4080
ggagctaggt aggattcagc taggttccta ctggggtacc cccctaccct accttaaggt 4140
gcgccaccct tcctccttcc gttttaatgg tagaataacc tataggttat taacctagtg 4200
gtggaatagg gtattgcagc tgggtatata cctataggta tatagaacct agaggaaggg 4260
aaccctatag tgtaatccct ccccccccta ccccccccctc ccttacggtt gcctgagccc 4320
atcccccacc ccagcacccc ggggtgacgt ggcaccccgc gtgccttact gacttgtcac 4380
ctttgcacat ttggtcagct gaccgatgct cgccacttcc tgggtcatga cctggcctgt 4440
gccttgtccc atggacaatg tccctccagc gtggtggctg cctttgggat gcatcacttt 4500
gagccactaa gcccccgttg ctcgccttgc ctgcctcacc atgacacact aagcccctgc 4560
taatccatga gccccgcctt taggaagcac cacgtcccgg ggacggaagc tggattttgg 4620
ccagtcttca attttgggga gtggttttgt gtgagccgga agttggcaat ggggtgaggg 4680
tggcgctggt taagctgacg acctcccaag gtctctcacc ctgggtacac aggtggggcg 4740
gcagcctcta actttggctg tggcctctat ttcctccctt tcctagccag ggccatgtgt 4800
tcctgcatgt ctacttgcct cctgtggtgg cagagcttgg ccctgggccc aacccccgcc 4860
ttgggagcct gtaggggcca acacccttgg tttgtttgtg ttcctgtttg ctggcaactt 4920
actggcagcc gagcagattc taatgggcgc ccgccttctt tctctcttgt tttattaata 4980
gaatctcagc caggacctat acctgagact tcaaagtctg gtcctgggtt ctgagacccc 5040
caagatttgt catgcacacc tgcacacctg ttggtattgg gtttctattc ttgagtgtga 5100

aagtttgtaa aaaaattcat aaaatgtcac taattcctct tacctgttta gggtattgtg 5160
caattcttca gcctgcctat tttcaatttg cctaaggtgg caatttaaga tgtggttaat 5220
taaccatttt cctgtctgac accactgcat gggcaaccgg gttccatggc acatttagag 5280
ataaacatag atgtcttgtc ttgctcatgt gcagaggagg gggtgttggt gtgcaatata 5340
gtttctggat tccaaattga gttgggggtg ctattttcac tatggaatta aattactgac 5400
attagacagt ggacaccggg ctatatgtgg ggatgtctgt ggcttgtcat ttcctcttag 5460
aaggtaatcc cccatcttaa cttcccttta aattgtgatg caagccctgg gttatttata 5520
gaatgattat ctaggtttga tagtctgaag gctgggcaga gaatgtttgt aatttttatt 5580
caccttcttt acccccccacg agtatccagt tctagaagat ctcctgatat cccgggctgc 5640
cattattccc ttgagtgtta tagcttcctc ttaacttaag caagagctcc aggatgttag 5700
ctttttggt ggggctggtt gtcaggaaga ggttccagtg ttgtccttta tttttagatg 5760
ttagctttgt gttaggttag tatgggctgg gtattcacta gtgaaggcaa ctaacacagt 5820
tagacgtgct agttgtgccc actggtgttt atccggtccc aaatgtcacc acagaacaca 5880
gggggctgga tttggcagca gcacttgtgc ttttgttgat ttttacccgt gtatcagagt 5940
gggggatgct agccaattta gcttcccctc cccttaacag ggggtctcgc ggggtgccaa 6000
ttgtcgcctg ccttcccccg cttccccttg ttaacttata gcatgatagg taggtcacct 6060
aacgtggaag cctggtgggt gatccttcct cggtagggag cgcttagggc tgttgagctc 6120
aacagcccca cctgggtaaa atgtatgttc taaagagtta cccaattata acaaaactgt 6180
tgtagggtaa cgaagacctg atggaagtgg tattgttgcc gttgaaagac gggtgtcctg 6240
gctcaagttc gcacttccta tacagtgtta aagccttgta tcggaagttt gggcttcgtc 6300
ccagtgtact cgataatgtc gactgctgcg aaaggtttgg accgtcttcc agtaggtgtt 6360
gggggtccca aatcacgagg ttaggcaggt gcacttggct ctttaggagg gacccttaag 6420
ccagacaatg tagtgccccc ttttttgca aattggcctt attattaatt tcttgttaac 6480
actaattctg ttctatgacc ctgtgttttt cagatgccgt tgaacgtgtc actgagctga 6540
atttggacgc agctacttga cctttgcccc cgtgcctcca gcgctgataa gtgctgcgtc 6600
cactttgtgt tacaggtggg ccaaacctcc agaatatcaa ttggtggggc cttggtgggc 6660
tgcataaggc agtaggtttg aggtgaccta cttggaccat gtggatccag tgtcctgatc 6720
ctggaccttg actatgaaac aattctaaaa aaatgcatca tagtccagtg tccagggaca 6780
gtgcactcgg aagtctcatc atctccgttt gtgtgtttag tgtggccagt acggccaccc 6840
ctgtgccacg ccctggcatg ctgctgacat ctggccgcca atttcagcgg gccctttcc 6900
cccttgttca ccccatagca agaagggtag gttacatggg tattttccca tcagcacctg 6960
actggccggt gcaattagag gagagggcaa caacgcaagg ctgttgtttt atttgggtta 7020
caagagctgc ggcggtcgat gggttcactg attacggttt cctagattgt acagatgaac 7080
tagaactgtc acaatctatg gggtcgtaga cagtgtgctt accagacttc catggaagat 7140
gtgaatttgc tgctagctat atgggtggtg ctatgggctc cctagggact catgtagtgg 7200
ggctttgtga tagctaatga atgtggcagc tgttgtttgt actggaccct gaattggaaa 7260
cagtaacttg gattctgtaa cacttcatgg gtcccgtagt gacaactatg ctgaatatct 7320
tgaatatggg aggaggggg ctttgggttc cattgtgtgc cctttcctgg ccaacgtgag 7380
ggtcctagtg ttatagggcg tggcagtttt cttgagggct aataacccgg gtgaggcggt 7440
tgtcacaggt gctagaccct ggagttgaac cagtaccact cggttacaaa gtcatggtct 7500
agtagttgtg accctgcaaa gctacgtggg gatgagcagc cagggacttt ggttggcaag 7560
cagacaggcg gcgcattgga accccagagg agtgtcccgg ggccacctct ttggttctgt 7620
acatattttg ttattgtaca taaccatgga gttggctgtg gtgcactcca tctggtaagg 7680
gggctggtgc ggacgcctgt gtttagtcta tgccaatgtt tacctgcctt gggttactat 7740
tccaaacgac cacacctttg aggacacctg gagccctgat cattctcggc ttttactgcc 7800
acctggcttc tgttgggtca gacagtttgg tgcgctagtt gtgtgcttag cagcaacgca 7860
caccaggctg actgccttag cagtgtggcc ctttattgtg gcatcctaag gagggattct 7920
ggagtgcctt tcgcgtgaag catgccctga gacgtactcg agttaggact taatcgctcc 7980
tgtgccgctg gatgagggag cgccaatttg tacatcctag ctctggccat agagttagcc 8040
caccttgtg tctccctttg gcctttgcgg tgccaatttc cggtggtttc ccttttccgc 8100
ccgtttatcc aatagcatgt aagagaggtt gcctagattt ggcaactttg agggaacgtt 8160
ccgtgtagct ggtgacctaa cacccgccca tcaccaccgg acagattctg aacttgtcct 8220
gtggtgtttg gtgtggtttt ggggtacgca ggagtacgtt ggaatgcttt ggagccgaga 8280
gggatgggcc cgcttgtgcg cttatgtgtt acacggtgcc aataaccggc ccggtgcggc 8340
tgccccgtga cccgtgggcc ttaccttcct ggccatcggg ggaccctggt gctagggtcc 8400
cttgtgttgc tttctgccat agggggggaaa gcatcgcctt cagaattggc tgctccgttg 8460
gaacatttga ggcctactgt atccgtgtcc tgacaacatt ccccgcaaac atgacatggg 8520
ttaatttaaa catgttttgt ttgcttggga atgctcttag ggcctggaag cttgtcattg 8580
gattcatcgt ttcctgaact acaggcgtag ggcctattgt agcaggcatg tcttcattcc 8640
tgcgtaccga atggcatgaa ggcacagcct gttaccattg gcacctttt tccatgtaaa 8700
cctccgtgat cctgggtcct ttggagactc aagtgtgaat ttgttttggt gttcggcgcc 8760
agggtatctc gacgttggaa tgtcaactca acttgggcac ctcgataacc ggctcgtggc 8820
tcgtacagac gattgtttgg ctctgtaact tgccagggac ggctgacgat gtgtttagtc 8880

tgccacttgc atccggcgct ttggttactc gggagactaa tgggggtgt ggtatggcac 8940
aggctggggg tgagtctggg gatgtccctg ggcgttgctg cagcccattc gccctctggg 9000
gatgagatgt tcaggggtgg ccggtaccct acgctgccga tttacataat ataaattgta 9060
aatgctgcag tagtagggat ctggacgcgc gacctgctac tcttcggaaa cgccaaccca 9120
ggagcgtcgc ctctggcccc atactcccgc catgcgactg ctcgcccct cccaggcctc 9180
cctggtgagc ccttgccgct ccccgcattc ctgctttcgg cgcccctgcg gatcccgatg 9240
acagcaggcc tttccttccc ccgttaatga aaagaatgac agtgaggttg tgacagaagg 9300
acagctttat tcagtttaca gagtgccctc ggaggctacg atattcccgt taaatgtctt 9360
gttgattctc tcaaaggtgg ggagggagga gctctccaca acaatgttcc ctggcagcgt 9420
gagcgcgcag ccctgccgtt ggatgtatct tctcatgatg gtgctgatag aggggtctcc 9480
ggcgtagatg aaaaaggcct gggccatgct ctggccggtc acgatcgtta tggggttgtt 9540
ggaaatgttc cggaccgtca gcttgagggt ctggcccggc ttccactcct gtgggtagac 9600
gtagaagacc gggttggagg agtgggacac gacaacggcc gtaatcttgg agctcagggg 9660
ggcctcgtag gtgttgttgt attccagctc cgtgatgaaa ttaggaggaa taatcacagg 9720
ggagccaaag tagcggatgt ctgtggattc cccgtcccag cgccagtggc tcttagggta 9780
ggggttgtaa cggaaggcaa taatcacatc atccaatagg gtcatgccca ccttgacgtt 9840
cagcgggccc tctcgtttca ggtccggcgt gtccacggag actcggacgt agcccttacc 9900
gcggcgtatg gcgtttaccg gacacacctt ccccgggaat gtgtgaatac gggcgtatga 9960
ctttagaaat gggggcgtgt gctgcgccag caggtaaggc aggcactcgt cctggctggt 10020
gacgggagag ccactgagga agatctgggg ctcgctggtg tttagcttgt ccccgctctg 10080
ggtgcaggag cgtgtcagct gaatgtcgct ctgcccgggc agaatctgca ggtagaggta 10140
ggggttcttg accaatctga tgggcacaat gtaccaggta aacttccctt tctctatgaa 10200
caggctgcgc ggattcagga cgcttagcac gatgtcctgg tcagagtgca taacgaagaa 10260
gggcttgagg aatacctcgt tgtcttccgc tccaaagaac aaaaacgcga ccgtaaagta 10320
gcggctgccg taggtggtcg tgttgaagga gaaagaaggt aacttgaagc tgagtatctg 10380
gcccaccgag gggcagggag gcagctcttg gcactgcgcg tccagctgca atacctgctt 10440
gttggtgacg cggacgtatg aggggaagat ctcgtacttc cacacgcctc tcatgaacga 10500
cgtgtctggt ttttcagtgg gccgcaggcg gcggaggctg ttcctgaacg acgagcgccg 10560
ggacgctagt gctgcatggg ctcctccggg gtaagcttcg gccatggccg gagctcgtcg 10620
acgggcaagg tgagagtcgg ggggcgggcg acggtgcggc cccaatacaa ctctccgctc 10680
gttagctggt agaatatccg cccggcgtct aggttgtcac ttcgctcggc cggccagaag 10740
agcgcaagtc caagtctggt gctggggccg atgtgcagcg gtttgtgccc gcagttgtag 10800
actgtcattt ttatgggcga gtgggcggtc cacacgcgcg ggcgcagcac ccattggtcg 10860
cacgccgcct cctggaatgt aaacccccag agagagggcg tgccgccctg gagatggccc 10920
tgtgccatca catgtatttc ctccttgggt ggaacaacgg cgtcgtgctc cgggtggagg 10980
gggaatagcg tccaggcatc tttcagggtc acgagaccgg ggtccatgct cagagaacag 11040
ccctcccggg cggtgggcgg cccgggctcc agcagaacgt cgcagaccca gccctcctcg 11100
gccctgtcca cctgtatgtc caggtgcacg gacccggagg ctgcgtctcg tgacatggcc 11160
aggcctggtg ccagccgacc acgtcccgtg tcccagccga ggccgcgcca gagcagagcc 11220
cgggactgac tcagggccac atcccctcgg cccgcggacg ccgcctcgcc agcccccggg 11280
ccttcatggg cccgctttct acctctctcc ggcaccccag cctggtcagc cgcagaggaa 11340
gcatgacctt ggggtgggac ggggcaggcg tgatcctggg cgcaatcttt gccgatcccc 11400
acaccttcac tccttgttag gttgatagaa tgtcggtacc acgccacggg gggcgggccc 11460
gcatagggaa aagccaggga gagcgatgtg ggcgaggatg ggctcaggcg gccccagaca 11520
cgcaatttgc ccccctgggc ggccgcagcc tgcccctcgg cggcccgtgc cccagctccg 11580
tcacgggggg cgcataggag gggtatatct aggatagccg cacctacaca aatgagacac 11640
agacacaggt cgtgaggatt taggcaacgc aggcttgtct ttatagttac aaacatggga 11700
gcgtgcacct ggaagatgca gctggggtag atctttacat ctttacaggg cgcagcggcc 11760
gccagacact gaagggcaga gttcacggcg ggcacctccc agagggagcc caccagcccg 11820
tacctggcca cggcc 11835

<210> 31
<211> 2337
<212> DNA
<213> Plasmodium falciparum

<400> 31
aaaaaagaaa attataaata aatatatata ttcgtgtaaa aataagtaga aaccacgtat 60
attataaatt acaattcatg atgagaaaat tagctatttt atctgtttct tccttttttat 120
ttgttgaggc cttattccag gaataccagt gctatggaag ttcgtcaaac acaagggttc 180
taaatgaatt aaattatgat aatgcaggca ctaatttata taatgaatta gaaatgaatt 240
attatgggaa acaggaaaat tggtatagtc ttaaaaaaaa tagtagatca cttggagaaa 300

atgatgatgg aaataataat aatggagata atggtcgtga aggtaaagat gaagataaaa 360
gagatggaaa taacgaagac aacgagaaat taaggaaacc aaaacataaa aaattaaagc 420
aaccagggga tggtaatcct gatccaaatg caaacccaaa tgtagatccc aatgccaacc 480
caaatgtaga tccaaatgca aacccaaatg tagatccaaa tgcaaaccca aatgcaaacc 540
caaatgcaaa cccaaatgca aacccaaatg caaacccaaa tgcaaaccca aatgcaaacc 600
caaatgcaaa cccaaatgca aaccccaatg caaatcctaa tgcaaatcct aatgcaaacc 660
caaatgcaaa tcctaatgca aacccaaatg caaacccaaa cgtagatcct aatgcaaatc 720
caaatgcaaa cccaaatgca aacccaaacg caaaccccaa tgcaaatcct aatgcaaacc 780
ccaatgcaaa tcctaatgca aatcctaatg ccaatccaaa tgcaaatcca aatgcaaacc 840
caaacgcaaa ccccaatgca aatcctaatg ccaatccaaa tgcaaatcca aatgcaaacc 900
caaatgcaaa cccaaatgca aaccccaatg caaatcctaa taaaaacaat caaggtaatg 960
gacaaggtca caatatgcca aatgacccaa accgaaatgt agatgaaaat gctaatgcca 1020
acaatgctgt aaaaaataat aataacgaag aaccaagtga taagcacata gaacaatatt 1080
taaagaaaat aaaaaattct atttcaactg aatggtcccc atgtagtgta acttgtggaa 1140
atggtattca agttagaata aagcctggct ctgctaataa acctaaagac gaattagatt 1200
atgaaaatga tattgaaaaa aaaatttgta aaatggaaaa atgttccagt gtgtttaatg 1260
tcgtaaatag ttcaatagga ttaataatgg tattatcctt cttgttcctt aattagataa 1320
agaacacatc ttagtttgag ttgtacaata tttataaaaa tatatactac ttttttttctt 1380
aattttcatt tttctttata ttttcctatt taatttattt ttttgtgaat atttaattac 1440
gtttgcgatt aattgtagaa atatatatgt atatactata tttatagaat gtgttattct 1500
caaaaacaac aacaaaaaaa aaaaaaaaaa aaaaaaaaag aaaaaaggat taaaagtaaa 1560
atagttataa atattttcaa aaatatttat aacacaaaaa atacttcgaa gttcatttaa 1620
cattttttgtt tatttattta tttatatatt tcatttttac gtatttatat tataaaatgg 1680
tgtatcttaa aaatagtgaa ctatatatat aaaatattaa tttaaaaaaa ttataacttt 1740
ctttttattt tctaaaataa cttaaaaatt atatgtttaa gaaaggggta aattataata 1800
tttgtataaa tatataaaca tagatatatt aaataaaata acaaatgtac tatatttgtg 1860
cataagacgt atacgcttta tataatacaa caatattaat tgtaataata tttgtggtag 1920
tgtgaacact aaaattgata ataatgatta taatacagaa gaaataaaaa atgaatccaa 1980
tataggattt acaacaaata ttcatgaagc aaaaataatt caagaaaaga catatggatt 2040
aataataaac gataaaataa agaaagaaga atatgatgat tgtaataata ataataataa 2100
taatattata atacagataa gagaagttgg acttaattat tttggagata ctctcgatga 2160
atcgaatcca tgtaatgatc ttacaggtat taatatatgg gaaagttgtc ttgtggctag 2220
tcgatggttt agcgatttat ctttacagaa tttttttcg aataaaaata ttttagaaat 2280
tggtgctggc agtggtttgg ctagtataat aatatttata tattctaata tttacaa 2337

<210> 32
<211> 729
<212> DNA
<213> Plasmodium falciparum

<400> 32
cgaaaaatat ttaattatct aaataaattt aattaaaaat ttttataaca tattttattt 60
aagattttat aataattaag ttttaatttc ttttgatcca aagtttttaa taattaaatt 120
tgtagatttt taatttattt aatatattca aaatgaaaat cttatcagta ttttttcttg 180
ctcttttctt tatcattttc aataaagaat ccttagccga aaaaacaaac aaaggaactg 240
gaagtggtgt tagcagcaaa aaaaaaaata aaaaaggatc aggtgaacca ttaatagatg 300
tacacgattt aatatctgat atgatcaaaa aagaagaaga acttgttgaa gttaacaaaa 360
gaaaatccaa atataaactt gccacttcag tacttgcagg tttattaggt gtagtatcca 420
ccgtattatt aggaggtgtt ggtttagtat tatacaatac tgaaaaagga agacacccat 480
tcaaaatagg atcaagcgac ccagctgata atgctaaccc agatgctgat tctgaatcca 540
atggagaacc aaatgcagac ccacaagtta cagctcaaga tgttacacca gagcaaccac 600
aaggtgacga caacaacctc gtaagtggcc ctgaacacta aacagctgta aacttttttg 660
ttaatgggtt tttttgaaac acgtgaaaat aatttttatt tatgattata ttatatatat 720
tgctatttt 729

<210> 33
<211> 594
<212> DNA
<213> Homo sapiens

<400> 33

atgggtgccc cgacgttgcc ccctgcctgg cagcctttc tcaaggacca ccgcatctct 60
acattcaaga actggccctt cttggagggc tgcgcctgca ccccggagcg gatggccgag 120
gctggcttca tccactgccc cactgagaac gagccagact tggcccagtg tttcttctgc 180
ttcaaggagc tggaaggctg ggagccagat gacgacccca tagaggaaca taaaaagcat 240
tcgtccggtt gcgctttcct ttctgtcaag aagcagtttg aagaattaac ccttggtgaa 300
tttttgaaac tggacagaga aagagccaag aacaaaattg agagagctct gttagcagaa 360
tgaaaaaatt ggaagccaga ttcagggagg gactggaagc aaaagaattt ctgttcgagg 420
aagagcctga tgtttgccag ggtctgttta actggacatg aagaggaagg ctctggactt 480
tcctccagga gtttcaggag aaaggcaaag gaaaccaaca ataagaagaa agaatttgag 540
gaaactgcgg agaaagtgcg ccgtgccatc gagcagctgg ctgccatgga ttga 594

<210> 34
<211> 1605
<212> DNA
<213> Homo sapiens

<400> 34
gtggcggcgg cggcatgggt gccccgacgt tgccccctgc ctggcagccc tttctcaagg 60
accaccgcat ctctacattc aagaactggc ccttcttgga gggctgcgcc tgcaccccgg 120
agcggatggc cgaggctggc ttcatccact gccccactga gaacgagcca gacttggccc 180
agtgtttctt ctgcttcaag gagctggaag gctgggagcc agatgacgac cccatagagg 240
aacataaaaa gcattcgtcc ggttgcgctt tcctttctgt caagaagcag tttgaagaat 300
taacccttgg tgaatttttg aaactggaca gagaaagagc caagaacaaa attgcaaagg 360
aaaccaacaa taagaagaaa gaatttgagg aaactgcgaa gaaagtgcgc cgtgccatcg 420
agcagctggc tgccatggat tgaggcctct ggccggagct gcctggtccc agagtggctg 480
caccacttcc agggtttatt ccctggtgcc accagccttc ctgtgggccc cttagcaatg 540
tcttaggaaa ggagatcaac attttcaaat tagatgtttc aactgtgctc ctgttttgtc 600
ttgaaagtgg caccagaggt gcttctgcct gtgcagcggg tgctgctggt aacagtggct 660
gcttctctct ctctctctct tttttggggg ctcatttttg ctgttttgat tcccgggctt 720
accaggtgag aagtgaggga ggaagaaggc agtgtccctt ttgctagagc tgacagcttt 780
gttcgcgtgg gcagagcctt ccacagtgaa tgtgtctgga cctcatgttg ttgaggctgt 840
cacagtcctg agtgtggact tggcaggtgc ctgttgaatc tgagctgcag gttccttatc 900
tgtcacacct gtgcctcctc agaggacagt ttttttgttg ttgtgttttt ttgttttttt 960
tttttggtag atgcatgact tgtgtgtgat gagagaatgg agacagagtc cctggctcct 1020
ctactgttta acaacatggc tttcttattt tgtttgaatt gttaattcac agaatagcac 1080
aaactacaat taaaactaag cacaaagcca ttctaagtca ttggggaaac ggggtgaact 1140
tcaggtggat gaggagacag aatagagtga taggaagcgt ctggcagata ctccttttgc 1200
cactgctgtg tgattagaca ggcccagtga gccgcggggc acatgctggc cgctcctccc 1260
tcagaaaaag gcagtggcct aaatcctttt taaatgactt ggctcgatgc tgtgggggac 1320
tggctgggct gctgcaggcc gtgtgtctgt cagcccaacc ttcacatctg tcacgttctc 1380
cacacggggg agagacgcag tccgcccagg tccccgcttt ctttggaggc agcagctccc 1440
gcagggctga agtctggcgt aagatgatgg atttgattcg ccctcctccc tgtcatagag 1500
ctgcagggtg gattgttaca gcttcgctgg aaacctctgg aggtcatctc ggctgttcct 1560
gagaaataaa aagcctgtca tttcaaacac aaaaaaaaaa aaaaa 1605

<210> 35
<211> 3107
<212> DNA
<213> Homo sapiens

<400> 35
atgatacagt ccaagcacct ggatgatgag tatgagagca gcgaggagga gagagagact 60
cccgcggtcc cacccacctg gagagcatca cagccctcat tgacggtgcg ggctcagttg 120
gcccctcggc ccccgatggc cccgaggtcc cagataccct caaggcacgt actgtgcctg 180
cccccccgca acgtgaccct tctgcaggag agggcaaata agttggtgaa atacctgatg 240
attaaggact acaagaagat ccccatcaag cgcgcagaca tgctgaagga tgtcatcaga 300
gaatatgatg aacatttccc tgagatcatt gaacgagcaa cgtacaccct ggaaaaggtg 360
ggtgcaggat gggagcagct ctgtgggggа agagcgggca tgggggtgcg gtgaccctgc 420
agcccctcaa ggcccagtct ctggagccat ctctcacctc tccgactctg agcttccact 480
gcactggcag tttgactcgt gcttcctgcc ctcggcttct gtctctcatg ctctctgagt 540
gtctcgccgt ctggccaggt gggtctcatc gcctctgcca gcgtcagctc ccacagcgaa 600

ggtcttccgt gtgctgtctt cttctgccct cgctcacgag tttggattcc ttgctgagga 660
gcagttctaa cccggaatca ctgtctgccg gcaggatgcc cagcatgggg tttggatctc 720
acactctgtt ttctccccca cgtagaagtt tgggatccac ctgaaggaga tcgacaagga 780
agaacacctg tatattcttg tctgcacacg ggactcctca gctcgcctcc ttggaaaaac 840
caaggacact cccaggctga gtctcctctt ggtgattctg ggcgtcatct tcatgaatgg 900
caaccgtgcc agcgaggctg tcctctggga ggcactacgc aagatgggac tgcgccctgg 960
gtatgattgg cctctccagc tcctcccctc ggtgctatcc tctggccaaa gaggtcctgg 1020
gattgcaata gcctggtggt ctggcgcaag ggcgtggggt gccctgggct cggtagagag 1080
caaaggatct caccagggcg gatggggaag cggtgctgga cgctgctcag ccctctctct 1140
gctctgtggc cccagatgac atctaagaga gacagtcaga gtcagggatt ccatcaaatc 1200
cctacctggg gcgcccctga ccaacagtcc tctggcctct gctgcatgcc caggcctcca 1260
cagcgactcc ccgggggctg ggaagtcata gtcatgctag ggagggcccc tgccaccgtc 1320
tctgctcatg gattcctttc cttgccctca gggtgaggca cccattcctc ggcgatctga 1380
ggaagctcat cacagatgac tttgtgaagc agaagtaagt atcacctgag ctaactgcgg 1440
ctctcactcg agcatccttt gtgtgctggt ctggctgaga aagcagttcc ctatcccaaa 1500
tcttcaactg gagggatggg tgcctctgac ctgggagtga gtggcagtgg ggggtatgcg 1560
agtgtgtggg gagccgaagg ccagggcggt cttgggaaaa gggagctcac gtcacctgag 1620
aacacggtgt ggggtgtgaa aacggccgcc atcaccttga gcacctgccc tgtagactga 1680
cacaagagtt ccccctggtt tacacctaag gaacccggag ctcagagagg agacgcctct 1740
gagcatggct cccagctggt aagggcctca gcccaactct cctgattttc aggccagggg 1800
ccaccctctc cccgtccctg gaggacttgc caacgcacag gcgcgcatgc acaccaacaa 1860
agggtcagga cttgaggagg atgcctggag cacgcttctc ctggctgact gtttcttcct 1920
ctccagtcgt ttcctctggt gggcctctcc agggctccgc cggggtgtgg ccaagaccct 1980
cgaggtgggg tgtgctcaga gcagggggcc tgaagaatgg ctcctctgtt tacaacacac 2040
ccaacaggaa gctggggtca tcgtgatgag gggcacaaac ttgtggcctc cctacagaca 2100
aatgccctac atgtggaccc cctgcacctc cgcatggctt ccggggagga ccaatggcaa 2160
aaggctttga aggcctcact tttgcaggca gaagtcctgg gagtgggttt gggaatgagt 2220
gaagggctgg aggggcagga cagtcctctt ccaggagctg agctgcggca tcgggttgag 2280
gaggggcccc ctggaaccca tccgttcagc aacaggtctg cttggctagc agcaaagttt 2340
actttcctct catgccaagg tacctggaat acaagaagat ccccaacagc aacccacctg 2400
agtatgaatt cctctggggc ctgcgagccc gccatgagac cagcaagatg agggtcctga 2460
gattcatcgc ccagaatcag aaccgagacc cccgggaatg gaaggctcat ttcttggagg 2520
ctgtggatga tgctttcaag acaatggatg tggatatggc cgaggaacat gccagggccc 2580
agatgagggc ccagatgaat atcggggatg aagcgctgat tggacggtgg agctgggatg 2640
acatacaagt cgagctcctg acctgggatg aggacggaga ttttggcgat gcctgggcca 2700
ggatcccctt tgctttctgg gccagatacc atcagtacat tctgaatagc aaccgtgcca 2760
acaggagggc cagcttcttc tcctggatcc agtaagagtt tcggcaccgt tgacgaactg 2820
cagcgatctt actggccaag ccagagcgcc tcctctcaga ttccttctcg acacagcacc 2880
ctaggcggct tcttcctgtc agtcggaggt ggcatgcaag atgaagctct ctttgctctt 2940
cccgctttca ttttgtgctt ttccttgtgt tttcatgttt tgggtatcag tgttacatta 3000
aagttgcaaa attaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3060
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaag 3107

<210> 36
<211> 2509
<212> DNA
<213> Homo sapiens

<400> 36
ccccactgtc tgccccgcag gctgtgcact gcccttccat catctggcag gcccccaaag 60
gtcagccccc ggtgccacac gagattccaa cgtcaatgga attccaggag gtgcagcaga 120
cacaggcgct ggcctggcag gcccagaagg cccccactca catctggcag cccctgcctg 180
cccaggaggc ccagaggcag gctcccccct tggtccagct ggagcagccc tttcagggag 240
ccccgccctc ccaaaaagcc gtgcaaatcc agctaccccc ccagcaggcc caggcatcgg 300
gtccgcaagc ggaggtgccc acactgccgc tccagccttc ctggcaggca ccgcctgcag 360
tcttgcaggc ccagcccgga cccccggtag cagcggcaaa ttttcccctg ggctccgcta 420
aatcattgat gactccatca ggagaatgca gggcctcttc tatagaccgc aggggctcct 480
ctaaagagcg caggacctcc tcgaaggagc gcagggcccc ttcaaaagac cgcatgatct 540
ttgctgccac cttctgtgct cccaaggcag tgtcagctgc gcgagcacac ctgccagctg 600
cctggaaaaa cctgcctgcc acaccggaga cctttgctcc ctcctcaagt gtcttcccag 660
ctacctccca gtttcagcct gcctctctga atgcctttaa aggcccctct gctgcctcag 720
agaccccaaa gtcactgcca tatgctctgc aggatccctt tgcctgtgta gaggccctgc 780

ctgcagttcc atgggtccca cagcccaata tgaatgcctc aaaggcatcg caggcagtgc 840
ccaccttcct gatggctaca gcagctgccc cccaggcaac tgccaccact caagaggcct 900
ccaagacctc cgtcgagccg ccacgccgct ccggcaaggc cacccggaag aagaagcatc 960
tggaagccca agaggacagc cgtggccaca cgctagcctt tcatgactgg cagggcccaa 1020
ggccctggga gaatctaaat ctgagtgact gggaggtcca aagccctatc caggtctcgg 1080
gtgactggga gcacccaaac acccccccgtg gcctgagtgg ttgggagggc cctagcacct 1140
ccaggatcct gagtggctgg gaagggccca gcgcatcctg ggccctgagt gcctgggagg 1200
gcccgagcac ctccagggcc ctgggtctct ctgaaagccc agggagctct ctgcccgtag 1260
ttgtgtctga ggtcgcaagt gtctctccgg gatccagtgc cacccaggat aattccaagg 1320
tggaggcaca gcccttgtct cccttggatg agagggcaaa tgcgttggtg cagttcctct 1380
tagtcaagga ccaagccaag gtgcctgtcc agcgctcgga gatggtgaaa gtcatcctcc 1440
gagagtataa agatgagtgc ttagatatca tcaaccgtgc caacaataag ctggagtgtg 1500
cctttggtta tcaattgaaa gaaattgata ccaaaaacca cgcctatatt atcatcaaca 1560
agctgggcta ccatacaggg aatttggtgg catcctattt agacaggccc aagtttggcc 1620
ttctgatggt ggtcttgagc ctcatcttta tgaaaggcaa ctgtgtcagg gaggatctga 1680
tctttaattt tctgttcaag ttagggttgg atgtccggga gacaaacggt ctctttggaa 1740
atactaagaa gctcatcacc gaagtgtttg tcaggcagaa gtacctagag tacaggcgaa 1800
tcccttacac tgagcccgca gagtatgagt tcctctgggg ccctcgagca ttcctggaaa 1860
ccagcaagat gcttgtcctg aggtttttgg ccaagctcca taagaaagat ccacagagct 1920
ggccattcca ttaccttgaa gcgctcgcag agtgtgagtg ggaagacaca gatgaggatg 1980
aacctgacac cggtgacagt gcccacggcc ccaccagcag gccccctccc cgctaatagg 2040
tgtagcagag atctcgctcc tgtgtttccc tggccagagg ccactgacag ggtgggggga 2100
catttttgtt cctggtgttt gtgttccagt tccacgagtg taagtttgga ttttcaactt 2160
ggtttcgtat ctgccaaagc tttgtacatt ttttatgtgg tgttgatttc aatcggctac 2220
tgttctgttc tgtattttgg catctgtgtt tttaagtgag atctgtggtt ctctgttttg 2280
tgttataatt gttatgtttt ggtatcagct ttgtgctggc tttgtgaaat gaattgagaa 2340
gctatccatc tcatttctgg tatagttcat gtagcattgt aatcggttgt tctttgaacg 2400
ttcaaatgac tcatcagtaa aaactgtcta cagagaagta aatatctata tctatatata 2460
taaatatact ttcagcataa aaaaaaaaaa aaaaaaaaaa aaaaaaaaa 2509

<210> 37
<211> 1782
<212> DNA
<213> Homo sapiens

<400> 37
agtgttgcaa ctgggcctgg catgtttcag tgtggtgtcc agcagtgtct cccactcctt 60
gtgaagtctg aggttgcaaa aggactgtga tcatatgaag atcatccagg agtacaactc 120
gaaattctca gaaaacagga ccttgatgtg agaggagcag gttcaggtaa acaaagggtg 180
ccacatctcc tgcctttctg ctcactttcc tgcctgtttt gcctgaccac agccatcatg 240
cctcggggtc agaagagtaa gctccgtgct cgtgagaaac gccgcaaggc gcgagaggag 300
acccagggtc tcaaggttgc tcacgccact gcagcagaga aagaggagtg cccctcctcc 360
tctcctgttt taggggatac tcccacaagc tcccctgctg ctggcattcc ccagaagcct 420
cagggagctc cacccaccac cactgctgct gcagctgtgt catgtaccga atctgacgaa 480
ggtgccaaat gccaaggtga ggaaaatgca agtttctccc aggccacaac atccactgag 540
agctcagtca aagatcctgt agcctgggag gcaggaatgc tgatgcactt cattctacgt 600
aagtataaaa tgagagagcc cattatgaag gcagatatgc tgaaggttgt tgatgaaaag 660
tacaaggatc acttcactga gatcctcaat ggagcctctc gccgcttgga gctcgtcttt 720
ggccttgatt tgaaggaaga caaccctagt ggccacacct acaccctcgt cagtaagcta 780
aacctcacca atgatggaaa cctgagcaat gattgggact ttcccaggaa tgggcttctg 840
atgcctctcc tgggtgtgat cttcttaaag ggcaactctg ccaccgagga agagatctgg 900
aaattcatga atgtgttggg agcctatgat ggagaggagc acttaatcta tggggaaccc 960
cgtaagttca tcacccaaga tctggtgcag gaaaaatatc tgaagtacga gcaggtgccc 1020
aacagtgatc ccccacgcta tcaattccta tggggtccga gagcctatgc tgaaaccacc 1080
aagatgaaag tcctcgagtt tttggccaag atgaatggtg ccactccccg tgacttccca 1140
tcccattatg aagaggcttt gagagatgag gaagagagag cccaagtccg atccagtgtt 1200
agagccaggc gtcgcactac tgccacgact tttagagcgc gttctagagc cccattcagc 1260
aggtcctccc accccatgtg agaactcagg cagattgttc actttgtttt tgtggcaaga 1320
tgccaacctt ttgaagtagt gagcagccaa gatatggcta gagagatcat catatatatc 1380
tcctttgtgt tcctgttaaa cattagtatc tttcaagtgt ttttctttta atagaatgtt 1440
tatttagagt tgggatctat gtctatgagc gacatggatc acacatttat tggtgctgcc 1500
agctttaagc ataagagttt tgatattcta tatttttcaa atccttgaat ctttttggg 1560

ttgaagaaga agaaagcata gctttagaat agagattttc tcagaaatgt gtgaaagaac 1620
ctcacacaac ataattggag tcttaaaata gaggaagagt aagcaaagca tgtcaagttt 1680
ttgtttttctg cattcagttt tgtttttgta aaatccaaag atacatacct ggttgttttt 1740
agccttttca agaatgcaga taaaataaat agtaataaat ta 1782

<210> 38
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 38
Gly Gly Gly Gly Ser
1 5

<210> 39
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 39
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> 40
<211> 70
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 40
acagacgtag cacagactat ggtattctcc agattaacag caggtattat gacggtagga 60
catgataggc 70

<210> 41
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 41
atagtttagc ggcccgcgat atctcactaa cactcattcc tgttgaagct 50